5-2013

TGF-B signaling pathway, ER-a and the heterogeneity of breast cancer risk among hispanic and non-hispanic white women.

Stephanie Denkhoff Boone
University of Louisville

TGF-B SIGNALING PATHWAY, ER-α AND THE HETEROGENEITY OF BREAST CANCER RISK AMONG HISPANIC AND NON-HISPANIC WHITE WOMEN

By

Stephanie Denkhoff Boone
B.S., Spalding University, 2006
MPH, University of Louisville, 2008

A Dissertation
Submitted to the Faculty of the
School of Public Health and Information Sciences
of the University of Louisville
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

Department of Epidemiology and Population Health
University of Louisville
Louisville, KY

May 2013

TGF-B SIGNALING PATHWAY, ER-α
AND THE HETEROGENEITY OF BREAST CANCER RISK
AMONG HISPANIC AND NON-HISPANIC WHITE WOMEN

By

Stephanie Denkhoff Boone
B.S., Spalding University, 2006
MPH, University of Louisville, 2008

A Dissertation Approved on

March 26, 2013

By the following Dissertation Committee:

_____
Kathy B. Baumgartner, PhD (Committee Chair)

_____
Richard N. Baumgartner, PhD

_____
Richard A. Kerber, PhD

_____
Frank D. Groves, MD

_____
Guy N. Brock, PhD

ACKNOWLEDGMENTS

I would like to extend a heartfelt thank you to my primary mentor and dissertation committee chair, Dr. Kathy Baumgartner. Your all-encompassing support, guidance, and passion for research have encouraged me to be more confident and independent and it is much appreciated. Special thanks to my secondary mentor, Dr. Richard Baumgartner, dissertation committee member and department chair, for your time, encouragement and willingness to critique my writing. I would also like to thank my dissertation committee members; Drs. Guy Brock, Frank Groves, and Richard Kerber, who made available their time and advice for me during the doctoral program and the dissertation. To my colleague and friend, Dr. Avonne Connor for her commitment in the lab and advice, I am very much appreciative. Special thanks to Christina Pinkston (the 'SAS guru'), for support and technical expertise in SAS programming. I would also like to acknowledge and thank the following people: Dr. Marty Slattery for manuscript editing, Jennifer Herrick for technical assistance with SAS, and Dr. Roger Wolff for guidance in the lab.

I give my heartfelt appreciation to my parents William Denkhoff and Karen Denkhoff, for the endless opportunities you have provided for me; you are the key to my success. I am grateful to all of my family and friends who have been supportive, loving, and encouraging. I owe my deepest gratitude to my husband, Matthew Boone; whose unwavering support and love has been my strength...I finally did it!

To end with, I would like to acknowledge funding sources throughout my four-year graduate research assistantship and dissertation: National Cancer Institute research contract '*Long-Term Quality of Life*' (R01-CA105266, Principal Investigator (PI): Dr. Kathy Baumgartner); Susan G. Komen Breast Cancer Disparities Epidemiology Research Training Program: (KG090926, PIs: Drs. Richard and Kathy Baumgartner); University of Louisville, School of Interdisciplinary and Graduate Studies, Sponsored Research Tuition Award and Doctoral Dissertation Completion Award. The Breast Cancer Health Disparities Study was funded by grant CA14002 from the National Cancer Institute to Dr. Marty Slattery at the University of Utah.

ABSTRACT

TGF-B SIGNALING PATHWAY, ER-α
AND THE HETEROGENEITY OF BREAST CANCER RISK
AMONG HISPANIC AND NON-HISPANIC WHITE WOMEN

Stephanie D. Boone

March 26, 2013

Many risk factors for breast cancer differ between race/ethnic groups. Few studies have included Hispanic women: a genetically admixed population that differs from other ethnic groups for breast cancer incidence, survival, and tumor phenotype. The objective of this study was to determine if genetic variation in *ERα* and *TGF-β* signaling genes (TGF-*β1, TGF-βRI, RUNX1, RUNX2, RUNX3)* is associated with breast cancer risk, and if these associations differ between Hispanic and non-Hispanic white women (NHW).

Data from The Breast Cancer Health Disparities (BCHD) study were used. BCHD is a multi-site consortium including two case-control studies within the U.S. and one in Mexico. A total of 3,524 cases (NHW=1,431; Hispanic=2,093) and 4,209 population-based controls (NHW=1,599; Hispanic=2,610) had available DNA. In-person interviews collected information on non-genetic risk factors. Single nucleotide polymorphisms (SNPs) in *TGF-β, RUNX* and *ERα* genes were determined using an Illumina platform and PCR. Associations with breast cancer risk were evaluated using multivariable logistic regression, adjusting for study site, age, and Native American genetic ancestry.

Associations with breast cancer phenotypes (ER/PR status) were also evaluated and a genetic risk score (GRS) was calculated to determine the cumulative effect of selected SNPs.

Two SNPs were significantly associated with breast cancer risk: *RUNX3* (rs906296 $OR_{CG/GG}$=1.15 95% CI 1.04-1.26) and *TGF-β1* (rs4803455 $OR_{CA/AA}$=0.89 95% CI 0.81-0.98). *RUNX3* (rs906296) was specifically associated with risk in pre-menopausal women (p=0.002) and in those with moderate to high Native American ancestry. There was a significant interaction between Native American ancestry and *RUNX1* (rs7279383, p=0.04). Four *RUNX* SNPs were associated with an increased risk of ER-/PR- (n=3) and ER-/PR+ (n=1) tumors. A GRS including 6 SNPs (range=0-10 alleles) across *ERα* and *TGF-β* signaling genes was positively associated with overall risk (per allele OR=1.14 95% CI 1.04-1.25), as well as ER+, but not ER- tumors.

These results suggest that genetic variation in these genes may explain the greater likelihood in Hispanic women for premenopausal, ER- breast cancer. This is also the first population-based observational study to evaluate crosstalk between *TGF-β* and *ERα* signaling pathways. The biological significance of these genes in breast cancer etiology is strongly supported and the results warrant confirmation in future studies.

TABLE OF CONTENTS

PAGE

TABLE OF CONTENTS

PAGE

TABLE OF CONTENTS

PAGE

LIST OF TABLES

TABLE PAGE

TABLE PAGE

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

INTRODUCTION

Breast cancer is the most frequently diagnosed cancer in women in the United States (U.S.) and Mexico in every major racial/ethnic group [1-4]. As a result of technological advances, the understanding of breast cancer has evolved to recognize that not all tumors are alike [5]. Epidemiological studies have attempted to evaluate the association of this tumor heterogeneity with lifestyle, ethnic, cultural, clinical, and biological factors in populations. Most studies have explored these associations in predominately non-Hispanic white women (NHW). Very few have included Hispanic women, a rapidly growing minority in the U.S., in which there are differences in risk of breast cancer and mortality, as well as risk factors associated with breast cancer in comparison to NHW women. Factors that explain these disparities remain to be established.

More than 70% of breast cancers may be characterized as "estrogen-dependent". The reproductive steroid hormone estrogen, which binds to estrogen receptors in the nucleus of the cell, is required for normal development as it stimulates cell proliferation in the breast epithelial tissue and activates target genes to produce specific growth-promoting proteins [6-7]. Estrogen is implicated in breast carcinogenesis because of the influence it has on growth and hormones that stimulate cell division, as well as its presence in most tumors diagnosed [7]. As a result, biological factors that influence the

activation or inactivation of estrogen-related carcinogenesis, such as susceptibility genes in hormone pathways, are primary targets for research and therapy. The extent to which genetic variation influences differences in breast cancer risk, between Hispanic and NHW, has not been established.

The primary objective of this dissertation was to evaluate differences between Hispanic and NHW women in the associations of selected genetic polymorphisms with breast cancer risk using data and DNA samples from The Breast Cancer Health Disparities (BCHD) study. The BCHD study is a multi-site consortium including 3 population-based case-control studies conducted within the U.S. and Mexico: the 4-Corner's Breast Cancer Study, the San Francisco Bay Area Breast Cancer Study, and the Mexico Breast Cancer Study [8]. Combined there are approximately 11,000 NHW and Hispanic/American Indian women diagnosed with a first primary breast cancer (*in-situ*/ invasive). The goal of this consortium is to evaluate the biological basis of ethnic-related health disparities, using genetic factors in the Convergence of Hormone, Inflammation and Energy Functioning (CHIEF) signaling pathway in combination with behavioral, social, and cultural factors [8]. The genes selected for study in this dissertation are limited to ones previously reported to influence the estrogen metabolism pathway and those that might explain differences in tumor phenotype: Estrogen Receptor-1 (*ESR1* or *ERα*), Runt-related transcription factor (*RUNX*) 1, 2, 3, Transforming growth factor-Beta 1 (*TGF-β1*), and Transforming growth factor-Beta Receptor I (*TGF-βRI*). Ultimately, data from this study could help illuminate genetic variation influencing ethnic disparities in susceptibility to breast cancer.

Specific Aims

Using data from a large, population-based case-control study, BCHD, which includes Hispanic and NHW women, the specific aims of this dissertation, are outlined below. Previous literature indicates a dual function for the *TGF-β* signaling genes (*TGF-β1, TGF-βRI, RUNX1, RUNX2*, and *RUNX3*) in breast carcinogenesis as promoters or suppressors depending on transcriptional influence of surrounding genes [9-10]. Corresponding hypotheses are therefore stated as alternative hypotheses without directionality of association based on this evidence.

(1) To determine if the risk of breast cancer is associated with individual single nucleotide polymorphism(s) (SNPs) in *TGF-β* signaling genes and *ERα* after adjusting for significant confounders.

(2) To determine if the statistical association of the risk of breast cancer and common variants in *TGF-β* signaling genes and *ERα* differ by proportion of Native American ancestry and menopausal status.

(3) To determine if genetic variants in *TGF-β* signaling genes and *ERα* differently affect the risk of breast cancer, as defined by Estrogen Receptor (ER) expression status, specifically ER+ and ER- tumors.

(4) To determine if the association of ER+ and ER- tumors with genetic variants within *TGF-β* signaling genes and *ERα* differ by proportion of Native American ancestry and menopausal status.

(5) To determine the statistical interactions between individual SNPs in *TGF-β* signaling genes and *ERα* and the association with risk of breast cancer.

(6) To evaluate the cumulative effect of *TGF-β* signaling genes and *ERα* with risk of breast cancer.

Hypotheses

H_1: Women who have the variant genotypes for *TGF-β* signaling genes will have an association with breast cancer compared to those who do not have the variant genotype, after adjustment for significant confounders. Women who have the variant genotypes for *ERα* gene will have a higher risk of breast cancer compared to those with the wild-type genotypes.

H_2: Women who have the variant genotypes for *TGF-β* signaling genes will have an association with ER- breast cancer compared to those who do not have the variant genotype, after adjustment for significant confounders. Variant genotypes for *ERα* gene will have a higher risk of ER- tumors compared to those with the wild-type genotypes.

H_3: The risk of breast cancer for women with the *TGF-β* signaling and *ERα* variant (s) will be modified by menopausal status and ethnicity as measured by proportion of Native American ancestry after adjustment for significant confounders.

H_4: The risk of ER- breast cancer for women with the *TGF-β* signaling and *ERα* variant (s) will be modified by a factor of estrogen exposure: menopausal status, after adjustment for significant confounders. The association will be modified by ethnicity as measured by proportion of Native American ancestry.

H_5: The proposed crosstalk between *TGF-β* signaling pathway and ER-signaling pathway is presented in mouse models. This will be demonstrated in this study population through statistical interactions among common genetic variants in *TGF-β* signaling genes and common variants in *ERα* gene when modeling the risk of breast cancer.

H_6: There will be an additive effect observed when combining risk alleles across pathways. The communication between signaling pathways is important when determining how genetic variation alters breast cancer risk. The present study will demonstrate the crosstalk with the cumulative effect of risk alleles and the association with breast cancer risk.

BACKGROUND

Racial/Ethnic Disparities in Breast Cancer Diagnosis and Outcomes

Approximately 23% of new cases of breast cancer (1.38 million) and 14% of the total cancer deaths (458,400) were ascertained in the U.S. in 2008 [3-4]. An estimated 226,870 new cases of invasive breast cancer will be diagnosed during 2012, with considerable variation across ethnic groups [1-2]. Cancer of the breast ranks second, behind lung, as a cause of death from cancer in women, with an expected 39,510 dying from this disease during 2012 [1]. Based on rates from 2007-2009, a woman living in the U.S. has a 12.4%, or a 1 in 8, lifetime risk of being diagnosed with breast cancer [2].

The age-adjusted incidence of breast cancer varies among race/ethnic groups by nearly 2-fold in the U.S. Data from the Surveillance Epidemiology and End Results (SEER) Program for the time period 2005-2009, indicate that NHW women have the highest incidence rate (127.3/100,000), followed by African American (121.2/100,000), Asian (94.5/100,000), Hispanic (92.7/100,000), and American Indian women (80.6/100,000) [2]. Cancer incidence rates for Hispanic women living in the U.S. have been available only since 1992. Although the breast cancer incidence rate for Hispanic women has been decreasing at a slower rate than NHW since 1997, 0.9% versus 1.5%, respectively; it is still 27% lower than that of NHW women. Conversely, evidence shows that breast cancer outcomes tend to be poorer among Hispanic women who are 20% more likely to die of this disease than NHW women diagnosed at a similar age and stage [11]. Because these population groups show disproportionate distributions in incidence and

mortality, epidemiological research has tried to explain these as ‘disparities’ in populations according to lifestyle, clinical, and biologically relevant characteristics [12]. The National Cancer Institute defines a cancer disparity as *‘a difference in the incidence, prevalence, mortality, and burden of cancer and related adverse health conditions that exist among specific population groups in the U.S.’* [13].

Reasons for the low incidence of breast cancer in Hispanic and Native American women are largely unknown. Due to the lack of detailed surveillance data in Mexico at this time, it is difficult to compare breast cancer rates across ethnic groups in two different countries, although, data suggest that the incidence is lower among Mexico-born Hispanic women compared to NHW women in the U.S. The BCHD study was designed to fill this knowledge gap by comparing Hispanic with NHW women living in the United States, as well as in Mexico, for breast cancer risk, taking European or Native American ancestry into account. Most women in Mexico are Hispanic by self-identification and are similar to those living in the Southwestern U.S. with respect to European genetic ancestry. Fejerman and colleagues found that the higher the proportion of European ancestry, the more positive the association was with risk of breast cancer in U.S. born Hispanic [14] and Mexico-born Hispanic [15] populations . They do note, however, that it is important to consider unmeasured confounding variables that could influence this effect and conclude that if there is an environmental element driving this association it must be common to U.S. Hispanic and Mexican-Hispanic born women [14-15].

In Mexico, reliable data for incidence and prevalence at the national level is not available due to the lack of cancer registries. Sources with available data, including National Institute for Statistics, Geography, and Information as well as the Ministry of

Health Database and the Mexican Social Security Institute, suggest that breast cancer is increasing steadily across the population, affecting all ages and socio-economic groups, with evidence of this increase exposed in hospital discharge rates associated with breast cancer, which increased by 80% for the time period 1986-2003 [16]. By 2006 breast cancer became the second leading cause of death among women aged 30-54 years in Mexico [16]. The significant burden of disease has challenged scientists to further research in basic science, epidemiology, clinical, and translational science to better understand breast cancer development and outcomes in diverse populations.

Established Risk Factors for Breast Cancer-'A Life Course Perspective'

Fluctuation in breast cancer incidence and outcomes can partially be explained by the changes in the risk factor profile of a population. Epidemiological studies have focused on determining the pattern or distribution of risk factors among various populations that may provide a biologically plausible basis for susceptibility, initiation and/or promotion of breast cancer. The Institute of Medicine (IOM) 2011 report, *'Breast Cancer and the Environment: A Life Course Approach'* [17] stressed the importance of a *'life course perspective'* as a tool for teasing out the significance of timing and interaction of exposures during a woman's life when evaluating breast cancer risk factors, as the breast is characterized by continuous change and hormonal influence. Individual behavioral choices, environmental exposures, and genes encompass a wide range of characteristics that influence the probability of developing breast cancer throughout critical periods of a woman's life. During these 'periods', ranging from *in-utero* through menopause, the breast transforms during gestation, menarche, pregnancy, breast-feeding, and menopause in response to varying levels of hormones, most commonly endogenous

or exogenous estrogen [17]. 'Breast tissue ageing', a term coined by Malcolm Pike, represents a mathematical model of the influence reproductive variables have on the risk of breast cancer, recognizing that the fluctuation in the rate of breast tissue aging is due to type of transformation the breasts undergo during critical periods [18]. An extension of Pike's mathematical model suggests the rate of breast tissue ageing is correlated with vulnerability to breast tissue damage [19]. Following this concept, researchers, including the IOM, suggest that these critical periods could be *'windows of susceptibility*' when the risk of breast cancer can be influenced depending on exposure to risk factors.

Estrogen Hormone

Estrogen is a steroid hormone that serves as a chemical messenger necessary for the normal development of the breast, as well as for regulating menstruation, reproduction, and maintaining the heart and healthy bones. The biosynthesis of estrogen occurs inside the cell in endocrine glands, specifically the ovaries and adrenal glands in women. Estrogen also can be produced via peripheral tissues (e.g. adipose tissues) stimulated by circulating steroid precursors, and can be formed in the liver, adrenal glands, breast, bone, and brain in women, where aromatase (an enzyme responsible for the conversion of steroid precursors to estrogen) is expressed [7, 20]. The influence of estrogen has profound effects on breast development throughout life; and is implicated in breast carcinogenesis due to its role of inducing cell division in the breast, its effect during the growth cycles when breast growth and development occurs; its interaction and effect on other hormones that stimulate breast cell division, and its consistent presence in the growth of estrogen-responsive tumors [7, 21].

The complexity of breast cancer etiology is not described exhaustively in this summary; rather it is evaluated with respect to established risk factors during a woman's reproductive life span that influence the level of exposure to the steroid hormone estrogen. Table 1 is a summary of the established risk factors for breast cancer by whether they cause an increase or decrease exposure to estrogen. The positive, negative, or inconsistent association of each risk factor with breast cancer is stated. NHW and Hispanic women can be characterized by different risk factor profiles according to epidemiological data and associations by these ethnic groups are presented. The overall magnitude of association is not provided as some studies have shown that one ethnic group may have a stronger association with a specific risk factor compared to another. Details for specified risk factors are included in subsequent sections.

Table 1: Established Risk factors for Breast Cancer by Estrogen Exposure and the Association among Hispanic and non-Hispanic white (NHW) Women[a, b]

Factor	Description	All	NHW	Hispanic
Increases Exposure to Estrogen				
Age at Menarche	Start of menarche (≤12 years)	+	+	+
Age at Menopause	Natural menopause (≥55 years)	+	+	+
Alcohol	Consumption (>1 drink/day)	+	+	+
Hormonal Therapy	Duration of estrogen use (≤ 2 years)	+	+	+
Oral Contraceptives	Current users	+	+	+
Obesity	Body mass index (≥30 kg/m²)	+	+	-
Decreases Exposure to Estrogen				
Parity	Trend for increasing number of full-term births	-	-	-
Age at First Full-term Birth	Early age at first full-term birth (≤20 years)	-	-	-
Breast Feeding	Duration of lactation	-	-	+/-
Physical Activity	Moderate to vigorous intensity over long-term period	-	-	-

[a] Symbol refers to association with risk of breast cancer: (+) positive or increase; (-) negative or decrease; and (+/-) inconsistent.
[b] References: [7, 17, 22-25]

Risk Factors That Increase Estrogen Exposure

Risk factors that increase exposure to estrogen can be either modifiable or non-modifiable. Risk factors that are non-modifiable include those that are part of the natural aging process and cannot be influenced by behavioral choices. These include: increasing age, early age at menarche, and late age at menopause. Cumulative exposure to endogenous and exogenous estrogen can affect the magnitude of lifetime risk of breast cancer. For instance, early age at menarche and late age at menopause, extends exposure to estrogen [26]. Research has shown that for every year menopause is delayed there is an additional 3-5% increase in risk, with a 30% increase in risk for age at menarche at >55 years compared to <45 years of age. Similarly, those who begin menarche at a later age (≥16 years) have a 15% decreased risk per year compared to those who begin at age 12 [27-28]. The direction of association has been found to be the same for both Hispanic and NHW women. Risk factors may also differ with respect to ER status and menopausal status, which are characterized by divergent levels of exposure to estrogen, and therefore may not have the same influence on risk. The risk for ER- tumors has been found to be higher for premenopausal women and for Hispanic women compared to NHW women, although the cause is unclear [29-30].

Modifiable risk factors that increase exposure to estrogen include: exposure to high levels of exogenous estrogen through use of post-menopausal hormone replacement therapy (HRT) or oral contraceptives (OC); obesity (body mass index (BMI) ≥30 kg/m²); and high alcohol consumption (>1 drink per day).

HRT. In a meta-analysis of nearly 4,000 cases conducted by the Collaborative Group on Hormonal Factors in Breast Cancer, post-menopausal HRT was found to

increase risk with duration of use. The risk was 40% (relative risk (RR) of 1.07 per year) higher for women who had taken HRT for five years compared to those not taking HRT for all tumor phenotypes combined. This estimate varied with dose of estrogen and progestin in HRT, and diminished each year after cessation [31]. The Multiethnic Cohort Study also found that HRT has the same effect on Hispanic (cases=257 of 11,792) and NHW (cases=419 of 13,659) women, although Hispanic women who were current users had a slightly higher risk (RR=1.36, 95% CI 5 1.20–1.54 *versus* RR= 1.26, 95% CI 5 1.17–1.37, respectively). Current HRT use was associated with ER+, but not ER- tumors [32].

OC use. The Collaborative Group also reported a 25% excess risk for current OC users based on a combined exposure. However, the effect diminished over time and after 10 years of cessation there was no association with risk [33]. Some studies have shown that there is a stronger association with current use and ER- tumors than ER+ tumors in premenopausal women [34-35].

Obesity. Obesity (BMI $\geq$30 kg/m^2) increases exposure to estrogen when aromatase, found in adipose tissue, initiates the conversion of precursor steroids, mainly adrenal androgens, to estrogens [7]. The association of obesity with breast cancer differs by menopausal status: post-menopausal obese women have an increased risk while premenopausal obese women have a decreased risk of breast cancer [21, 25]. Some studies suggest this association may differ by ethnicity: in NHW obese women the risk is divergent by menopausal status, in contrast, for Hispanic obese women menopausal status does not affect the association of obesity and breast cancer [36]. The 4-Corners Breast Cancer Study reported that among premenopausal NHW women, obesity was

associated with a higher risk of ER- tumors (Odds Ratio (OR) = 2.47; 95% CI: 1.08-5.67) but was inversely associated with risk of ER- cancers among Hispanic women (OR = 0.29; 95% CI: 0.13-0.66) [37]

Alcohol consumption. While the etiology of alcohol and breast cancer is still under evaluation, the associated risk is hypothesized to be a result of an interaction between alcohol and estrogen metabolism or the metabolism of alcohol by-products, which cause DNA damage [38]. Alcohol is recognized to be associated with a moderate increase in breast cancer, and the range of intake appears to have a linear relationship with risk. In 2006, Key and colleagues [22] conducted a meta-analysis of 98 studies, which showed that each additional 10 grams of alcohol/day was associated with a 10% increased risk (95% CI: 5–15%). The National Institutes of Health - AARP Diet and Health Study (1995–2003) evaluated the alcohol and breast ER status association and found alcohol (5-30 grams/day compared to non-drinkers) was associated with ER+ but not ER- status [39].

Risk Factors That Decrease Estrogen Exposure

Risk factors that decrease the exposure to estrogen include: parity, age at first full-term birth (FFTB), breastfeeding, and physical activity.

Parity. Parity reduces the number of menstrual cycles and cumulative exposure to estrogens and induces the full differentiation of breast epithelium, making it more resistant to carcinogenic transformation. After the first half of pregnancy and during lactation, estrogens do not affect breast growth and differentiation. Epidemiological evidence suggests that women who have had a least one full-term birth have a 25%

reduced risk compared to nulliparous women, with protection increasing by as much as 11% with the birth of each child [23, 40-41].

Age at FFTB. Early age at FFTB is associated with a reduced risk of breast cancer and appears to be independent of the total number of births. Women who have their FFTB before age 20 have half the risk of developing breast cancer compared to women with a FFTB at age 30 [42]. It has been reported that this association appears to be associated with an increased risk of developing ER+ disease [41].

Lactation. Breastfeeding for long duration (at least one year) is reported to be associated with a decreased risk of breast cancer, as much as 25% in general as well as for both ER+ and ER- tumors [41]. The New Mexico Women's Health Study assessed reproductive factors among Hispanic and NHW women and reported an inverse association of higher parity with breast cancer risk for NHW, but not Hispanic women ($p < 0.008$). Longer lactation was associated with reduced risk in premenopausal but not post-menopausal Hispanic women similar to the associations in NHW women [43].

Physical activity. Physical activity decreases risk of breast cancer when measured at a consistent level of intensity over a long-term period [24]. The biological mechanisms may include reduced estrogen levels, decreased factors involved in inflammatory and immune responses, and maintenance of a healthy bodyweight. Most studies indicate that at least 30 minutes of moderate to vigorous level intensity per day is associated with a decreased risk; and a dose-response relationship is observed with level of intensity. However the magnitude of effect across studies ranges from a 20-80% risk reduction [24]. A systematic review of 19 cohort and 29 case-control studies reported through 2006 reported that the effect is stronger in post-menopausal women [24].

Interestingly, this effect has been found to be similar in post-menopausal women regardless of ethnicity. Findings for pre-menopausal women indicated a reduction in risk for Hispanic women only [44].

Risk Factors Not Associated with Estrogen Exposure

Other factors not related directly to estrogen exposure that are important to consider are family history and socioeconomic status (SES). Having a first degree family history of breast cancer is consistently associated with a 2-fold increase in risk compared to those with no family history, and risk increases for each additional relative affected [26]. The 4-Corner's Breast Cancer Study reported that risk for an ER- tumor was higher for Hispanic women with a family history, whereas risk for an ER+ tumor was higher for NHW women [45].

Several studies have shown that high SES is associated with increased risk of breast cancer across all ethnic groups, however the strength of the association is variable; it is stronger among NHW than Hispanic women at the highest SES level [46-48]. SES is a measure of individual or family attributes that contribute to a standard of living and common indicators are annual income, educational attainment, or occupational status. These factors are correlated with other risk factors such as parity, HRT, and alcohol use [47]. Low SES is reported to be associated with more aggressive ER- tumors in Hispanic but not in NHW women [49].

Use of Risk Factors in Statistical Analysis

The determination of risk factors and their contribution to breast cancer etiology is complicated. The RR and OR estimate the strength and direction of the statistical association of a risk factor with breast cancer, but do not necessarily offer insight into

cause or mechanism or whether the impact is necessarily associated with a large proportion of cases [6]. However, if results are consistent across several studies with differing designs, and cannot be attributed to bias, confounding, or chance, then a risk factor may be recognized to be a potential etiologic factor [6]. These established risk factors can then be used as a means to evaluate other predictors (i.e. genes), and be used as confounders or effect modifiers of the independent variable (also known as the main predictor) and dependent variable (also known as the outcome) when testing for an association. Most risk factors can be viewed as confounders, when they are associated with both main predictor and outcome, and not part of the causal pathway. They will be accounted for by adjustment in statistical modeling [50]. An effect modifier is a variable that differentially (positively and negatively) modifies the association of a risk factor with an outcome. The effect modification is detected when stratification reveals the effect estimates diverge across strata [50].

Most of the established risk factors discussed above were assessed as potential confounders in this dissertation (current age, age at menarche, age at menopause, alcohol consumption, HRT, OC use, obesity, parity, age at FFTB, physical activity, family history, and SES). Some risk factors such as menopausal status, and ethnicity were also evaluated as potential effect modifiers.

Breast Cancer Heterogeneity

When evaluating the biological features of breast cancer it is important to understand that it was once studied as a single disease and is now widely acknowledged as several distinct phenotypic subtypes that vary in etiology, gene expression, clinical features, and response to treatments, prognosis and outcome [5-6]. Most breast tumors

are associated with tumor markers that represent over-expression of specific reproductive hormone receptors and proteins, and their associated genes, that are involved in both initiation and promotion of carcinogenesis within the breast tissue [6]. However, some breast tumors do not express these markers and are believed to develop from a completely different biological mechanism. The most widely studied tumor marker is ER, which is a product of the *ERα* gene. The degree of tumor *ERα* expression (positivity) has been used for some time to classify breast tumors into subtypes believed to predict patient response to treatment, risk of recurrence and survival [6, 35, 51]. The Progesterone Receptor (PR), and Human Epidermal Growth Factor Receptor II (HER2), encoded by *PGR* and *ERBB2* genes, respectively, is used increasingly in combination with ER status to classify tumors into subtypes.

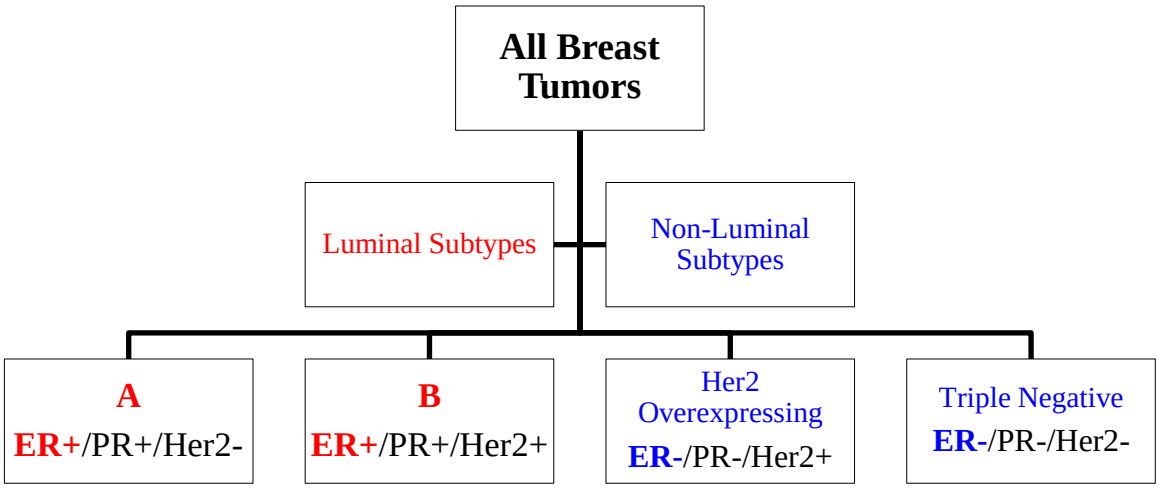
The ability to establish tumor heterogeneity can be attributed to advances in molecular and cell biology and lab techniques involving DNA sequencing potential. Classification of breast tumors is part of the routine diagnostic tests and is commonly performed using two methods: immunohistochemistry (IHC) and gene expression profiles (GEP) [52-53]. IHC is performed in individual samples of formalin fixed, paraffin-embedded breast tumor tissue blocks (from core-needle biopsy or resections). The technique utilizes a Food and Drug Administration approved IHC kit that includes monoclonal antibodies that can bind to specific antigens on proteins in the cell, in this case those expressed by receptors such as ER/PR/HER2, to test for a proportion of cells stained positive as well as the intensity of staining [52-54]. In clinical practice, a patient with a proportion of cells ≥1% staining for ER, regardless of staining intensity, will usually benefit from endocrine therapy, and the American Society of Clinical Oncology

and College of American Pathologists (ASCO/CAP) has set this as the standard cut-off value for ER positivity [52]. According to ASCO/CAP, the cutoff value for PR positivity status is the same as ER, while HER2 positivity is defined as a score of 3+, which corresponds to ≥30% of cells positively stained [52, 55]. These classifications are most commonly reached through a semi-quantitative method in which the number of positively stained cells are estimated on a slide cut from the tumor block by a trained pathologist [53]. Four intrinsic breast tumor phenotypes have emerged according to ER/PR/HER2 receptor status that corresponds to those established by gene expression (Figure 1). They include: luminal A tumors (ER+ and/or PR+/HER2-), accounting for 65-70%, characterized by over-expression of ER, PR, and luminal-specific genes; luminal B tumors (ER+ and/or PR+/HER2+), accounting for 7-12%, and having an expression pattern similar to luminal A but with HER2 overexpressed; HER2-overexpressing tumors (ER-/PR-/HER2+), accounting for 6-10%, which do not express ER/PR, but are found to have high levels of HER2 expression; and triple negative, (ER-/PR-/HER2-) which account for approximately 6-10% of the distribution [5, 56]. A full description of all breast tumor phenotypes is beyond the scope of this review as the focus is on ERα specifically.

While data have emerged describing incidence trends for breast cancer tumor phenotypes in different populations, the etiology associated with this heterogeneity is largely unknown and is the subject of growing epidemiologic research. It is well established that family history is a strong predictor when determining risk, supporting the need to explore the contribution of genetic variation to breast cancer, and the influence

this variation may have over the genes that represent different tumor phenotypes and heterogeneity.

Figure 1: Classification of Breast Cancer Subtypes by Hormone Receptor Status



Adapted from Blows, F.M., et al., *Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies.* PLoS Med, 2010. **7**(5)***[56]***.

Genetic Variation and Predisposition to Breast Cancer

The human genome contains over 20,000 genes that encode proteins. A very small portion of the genome, approximately 1.5%, contains coding regions while the rest are non-coding [57]. All cancers involve the failure of genes to control cell growth and division. Although a small proportion of breast cancers are strongly hereditary; it has been of interest to evaluate genetic variation in potential susceptibility genotypes and the association with breast cancer, as only a handful of genes to date explain the association [58]. There are different forms of genetic variation that are commonly described by the influence they have in breast carcinogenesis.

High Penetrance Mutations

High penetrance mutations present a high risk but are very rare in populations. *BRCA1/BRCA2* [59], and *TP53* [60], are a few genes with these types of mutations. They have been associated with relative risks that are 10 to 20-fold higher than risk for non-carriers of these mutations; however they only account for approximately 5% of inherited breast cancers.

Moderate Penetrance Variants

Moderate penetrance variants are uncommon with a minor allele frequency ranging from 0.005-0.01. Variants found on some of the following genes that are involved in DNA repair mechanisms include: *CHEK2* [61], *ATM* [62], and *BRIP1* [63]. The risk is 2 to 4-fold higher than risk for noncarriers of the variant but only account for about 3% of inherited breast cancers.

Low-Penetrance Variants

Low-penetrance variants include the most common forms of genetic variation and are often SNPs. These variants may increase risk <2-fold in some populations but have no effect in others [58, 64]. Large scale collaborative efforts such as International HapMap Project and Genome Wide Association Studies (GWAS) have identified approximately 12 susceptibility SNPs for breast cancer. Some of these SNPs are located on *FGFR2, TOX3, MAP3K1,* and *LSP1* genes [65]. GWAS have been conducted primarily in women of European background. Some of these results have been replicated in other populations, including a subset of this study population to be examined in this dissertation (4-Corner's Breast Cancer Study) [66-68]. Most variants are associated with ER+ tumors, although several population-based studies have identified a small number

associated with ER- tumors that appear to increase risk by only 3-14% [67]. Many are not considered to be causal variants but are thought to be in linkage disequilibrium with a neighboring functional variant [58].

Population Substructure in Genetic Variation

When assessing genetic variation, it is important to consider the population substructure. The phenomenon 'population substructure' is an effect seen in populations that could potentially differ genetically [69]. This substructure may be reflected in incidence of disease, proportions of cases and controls, or distribution of allele or genotype frequencies between groups and individuals. Undetected population substructure can create a false positive association due to the bias of missed genetic heritage or ancestry [69]. One successful approach to controlling this issue is assessing genetic ancestry, also known as admixture, using a set of genetic markers. The present study includes Hispanic women from the Southwest U.S. and Mexico who have Native American and European genetic heritages [8]. An estimate for the proportion of ancestry is used as a tool to control for potential confounding due to linkage disequilibrium within a genetically admixed population (See Methods-*Genotyping-All Study Sites*).

Genes of Interest

In keeping with the goal of the BCHD consortium, this dissertation study has evaluated genes that are related to the *CHIEF* pathway. The acronym '*CHIEF*' was first described by Slattery and colleagues [70] using colorectal cancer (CRC) as an example. They describe the *CHIEF* pathway, including loci pertaining to inflammation, hormones and energy that are interrelated and have an effect on risk, to demonstrate the complexity of CRC etiology. They further hypothesize that this pathway is also related to the

etiology of other cancers, specifically breast and prostate. In breast cancer, hormone-related loci are key determinates for risk and also have an effect on tumor promotion. Inflammatory loci are also involved in tumor growth when influenced by interaction with epithelial and vascular endothelial cells, which are partly regulated by genes in the *TGF-β* signaling pathway. Energy-related factors include those related to insulin-signaling such as energy balance, physical activity, obesity, and diabetes. Some of these factors, like obesity, are also involved in the inflammatory pathway. In this dissertation study, genetic variation in the *ERα* and *TGF-β* signaling genes represents the interaction between the hormonal and inflammatory portions of the *CHIEF* pathway and are described below.

ERα Signaling Pathway in Breast Cancer

The *ESR1* (*ERα*) gene is part of the nuclear receptor subfamily 3 group A member 1 (NR3A1) located on chromosome 6q24-q27, from positions 152,011,631-152,424,409; approximately 300 kb long and includes 8 exons or 'protein-coding regions' [71]. This gene encodes for the nuclear receptor *ERα*, a ligand-binding transcriptional factor that contains 5 structural functional domains (A-F) that are responsible for binding to steroid hormones and DNA, nuclear localization, and transcription activation [72].

Biological studies using mouse models have described two signaling mechanisms for *ERα*. The classic *ERα* signaling pathway involves the ligand-binding domain (LBD), which is encoded within a region of about 300 amino acids, binding with estrogen that passively diffuses into the cell [7]. The binding causes a conformational change, called dimerization, in *ERα*, which then translocates into the cell nucleus. The *ERα*-estrogen complex contains two zinc-fingers that allow it to bind to certain DNA sequences in the promoter region of the target genes known as estrogen response elements (ERE). The

binding regulates gene transcription and is the primary mechanism by which estrogen either activates or represses gene expression, or simply turns it 'on' or 'off' [73-76]. The second related mechanism, "membrane-initiated signaling", has been evaluated more recently [77], and involves activation of *ERα* on cell surfaces, signaling a cascade of events that regulate gene expression.

Through these inter-related mechanism(s), *ERα* mediates effects of estrogen on the expression of genes that regulate cellular development, reproduction, proliferation, and homeostasis. Because the binding of estrogen to *ERα* plays an extensive role in many physiological processes it has been linked with the development of several cancers and diseases besides breast cancer. These include cancers of the endometrium, ovaries, colon, and prostate and diseases such as cardiovascular, fibroids, endometriosis, osteoporosis, insulin resistance and obesity [72, 78].

There have been two proposed biological mechanisms for estrogen-related carcinogenesis in the breast. First, cell proliferation is stimulated in the mammary tissue when estrogen binds to *ERα,* causing an increase in target cell numbers, cell division, and DNA synthesis. This increases the risk of replication errors that could disturb normal cellular function. Second, the metabolism of estrogen directs the production of genotoxic by-products that could cause damage to DNA, resulting in point mutations [7].

The expression of proteins transcribed by genes affected by the *ERα*-estrogen binding is complex and highly regulated, suggesting that they may serve different functions depending on the stage of tumor development or progression [78].

Heterogeneity of ERα in Population-based Studies. *ERα* is overexpressed in approximately 75% of breast tumors and roughly corresponds to the "luminal" subtype

established by gene expression [6]. Endocrine therapy (i.e. tamoxifen and/or aromatase inhibitors) is often prescribed for ER+ luminal tumors in addition to surgery because they are considered "estrogen-dependent". ER- tumors are non-luminal and are associated with other tumor characteristics such as later stage, poorly differentiated grade, and larger tumor size, which result in a poor prognosis because they respond differently, if at all, to endocrine therapy. Women with this tumor phenotype are usually prescribed various types of non-endocrine chemotherapy. In addition, luminal subtypes are more likely than non-luminal to present at diagnosis as smaller, well-differentiated tumors at an earlier stage. Established risk factors for breast cancer are mostly associated with ER+ tumors. Little is known about the etiology of ER- breast tumors or what treatment appears to be most beneficial. Due to the fact that there is targeted therapy for ER+ tumors, women diagnosed with this subtype have a 90% 5-year survival rate compared to 50% in women diagnosed with ER- tumors, with most breast cancer-specific deaths occurring within the first 5 years following diagnosis [29, 35, 79-81]. With regard to ethnic disparity, while Hispanic women have a lower incidence rate of breast cancer than NHW women, they are at increased risk for development of tumors with less favorable clinical characteristics such as later stage, ER- status, and a poorly differentiated grade. The reasons for these differences are not well established [82-86]. A large number of SNPs in the *ERα* gene have been evaluated in population-based studies and observed to have an association with breast cancer; however, very few studies have been able to assess Hispanic women.

TGF-β Signaling Pathway in Breast Cancer

Genes in the *TGF-B* signaling pathway have been implicated in the development and progression of breast, colon, gastric, pancreatic, and ovarian cancer [87-88]. There is

increasing evidence that the *TGF-β* signaling pathway plays a dual role in cancer, acting as a tumor suppressor and promoter through a *SMAD*-mediated process [10]. To simplify the nomenclature, *SMAD* is used to describe homologs of both the Drosophila protein, mothers against decapentaplegic (*MAD*) and the Caenorhabditis elegans protein *SMA* (from gene sma for small body size) [89]. The *SMAD* family includes intracellular proteins that transduce signals given by *TGF-β* from the cell membrane to the nucleus providing instructions of protein production. In the nucleus, the *SMAD* complexes bind to specific areas of DNA, and act as transcription factors where they control the activity of particular genes [89]. The following *SMADs* are associated with *TGF-β* signal transduction: receptor regulated (*R-SMAD*) *SMAD2* and *SMAD3*; common mediator (*co-SMAD*) *SMAD4*; and inhibitory (*I-SMAD*) *SMAD6* and *SMAD7* [89]. Their specific roles in the signaling pathway are described below.

TGF-β1 and TGF-βRI. *TGF-β1* is part of the *TGF-β* family of multifunctional cytokines, and is located on chromosome 19q13.1 from positions 41,836,811 to 41,859,830; approximately 25 kb long, encoding 390 amino acids [90]. *TGF-β1* is an abundant peptide expressed in many types of cells. It functions in regulating biological processes such as cell proliferation, differentiation, adhesion, migration, and survival. A general biological mechanism for *TGF-β* signaling has been described [91]. Upon activation, the *TGF-β* ligand is cleaved from *TGF-β1*, binding to *TGF-βRII*, a high affinity cell surface receptor. It then recruits and phosphorylates *TGF-βRI* activating its kinase function. *TGF-βRI* is a cell surface receptor in the serine/threonine kinase family located on chromosome 9q22.33 from positions 101,867,411 to 101,916,473; approximately 49 kb long, containing 9 exons [92]. *TGF-βRI* initiates signal transduction into the intracellular

matrix with its protein kinase activity, phosphorylating *R-SMAD2* and *R-SMAD3*. These phosphorylated *R-SMADs* are then translocated to the nucleus, forming a complex with *Co-SMAD4*, and work to either activate or repress target genes regulating cell proliferation through interaction with other transcription factors. *SMAD6* and *SMAD7* inhibit *TGF-β* signaling by blocking *TGF-βRI* from phosphorylating *R-SMAD2* or *R-SMAD3* [87, 91, 93].

In normal mammary cells, *TGF-β1* has been found to have an anti-proliferative effect on epithelial and endothelial mammary cells by acting as a tumor suppressor by down-regulating cell growth, differentiation and apoptosis [87-88]. This is done by down-regulating components of the cell cycle such as proto-oncogene c-myc or cyclin-dependent kinases (CDKs) and up-regulating CDK inhibitors [87-88, 93]. Some researchers have found evidence in mouse models that increased levels of *TGF-β1* in serum strengthen tumor suppressor activity, reducing risk of breast cancer [94]. Immune cells, including B-Cell, T-Cell and macrophages, secrete *TGF-β1*, which negatively regulates their proliferation, differentiation and activation by other cytokines. This makes *TGF-β1* an effective immunosuppressor, and disruption of the signaling pathway is linked to autoimmunity, inflammation and cancer [95].

In most tumor cells, genetic variation in key members of the pathway can cause resistance to the growth inhibitory effects of *TGF-β* signaling [96-97]. Exact mechanisms for resistance remain unknown, although researchers have hypothesized, through evidence in gastric, pancreatic, and colon cancer studies, that there may be decreased expression of receptors on the cell surface or increased expression *I-SMAD6* or *I-SMAD7* in the extracellular matrix, inhibiting the signaling function [87]. Some

researchers suggest that reduced expression or inactivation of *TGF-β* signaling could be caused by oncoproteins such as p53 [98] or decreased expression of other tumor suppressors that regulate the pathway such as *RUNX3* [99].

There is also evidence that when *TGF-β1* and *TGF-βRI* are overexpressed following tumor initiation, they further promote angiogenesis or the development of new blood vessels from pre-existing ones, a condition that is necessary for tumors to grow larger than 1-2 cm [87]. One possible explanation is that *TGF-β* induces the expression of vascular endothelial growth factor (*VEGF*), which then directly promotes angiogenesis, leading to tumor progression and metastasis [100].

In summary, *TGF-β1* and *TGF-βRI* have been found to act at two different stages of carcinogenesis. First, at cancer initiation, when they can act as tumor suppressors until a disruption causes resistance to its growth-inhibitory effects due to the loss or mutation of members of the pathway. Second, during cancer progression, when there are tumor promoting effects, including enhanced motility, adhesion and angiogenesis in response to increased expression of *TGF-β1* by the cancer cells themselves [87]. SNPs on these genes have not been implicated in risk of breast cancer in GWAS studies; however, an increasing number of epidemiological studies have suggested that genetic variants affecting *TGF-β* production and/or signaling may be related to the overall risk of breast cancer. Several SNPs and/or mutations in *TGF-β1* and *TGF-βRI* have been associated with increased breast cancer risk and ER+ tumors, although results from several studies involving multi-ethnic or small sample sizes have been inconsistent [96, 101-103].

RUNX1, RUNX 2, and RUNX3 Transcription Factors. The *RUNX* family includes 3 genes: *RUNX1*, *RUNX2*, and *RUNX3*. They bind to DNA through the 128 amino acid

Runt domain (α-subunit), and share a common heterodimeric binding cofactor, called a core binding factor-beta (*CBF-β*). The *RUNX* family members have been found to play an important tissue-specific role in determining the fate of cells during differentiation and growth and there is increasing evidence that the loss of function of these genes is involved in carcinogenesis [104-106]. Although widely expressed, the *RUNX* family members are regulators of tissue-specific expression and there is suggestion that one cannot compensate for the other if there is a loss of function during development, which is evident in the mouse knockout phenotypes (*RUNX* deficiency) [107]. Because of the regulatory role of *RUNX* proteins, there is physical interaction with *R-SMADs* and stimulation by the *TGF- β* signaling pathways, which mediate these cellular functions [108]. *RUNX* proteins are described as downstream effectors of *TGF- β* signaling and have the ability to stimulate growth regulation by making target cells sensitive to the effects of *TGF- β* family members. In turn, *TGF-β* genes can activate *RUNX* genes at the transcription and post-transcriptional levels [9, 109].

RUNX1 is located on chromosome 21q22.3 from positions 36,160,097 to 36,421,594, and includes 11 exons and 453 amino acids and spans 260 kb[110]. In normal cells, this gene is involved in regulation of hematopoiesis. It is well-known for being the site of translocations in acute leukemia (AML and ALL), where they act as transcriptional repressors by hindering transcription of the wild-type allele and is often found amplified in these cancer cells [111-113]. *RUNX1* protein has been found at high levels in normal luminal and basal cells of breast epithelium, whereas expression is low or deficient in breast tumors [114]. Interestingly, *RUNX1* has been suggested to be a regulator of breast tissue development interacting with a family of transcription factors

called *FOXO* proteins [107, 115]. In-vitro studies indicate that there is an inverse relationship between *FOXO* and *RUNX1*, where the loss of *RUNX1* expression, causing oxidative stress on cells, is compensated for and stabilized by *FOXO*. The study reported that the *FOXO* expression is essential for breast cancers, specifically triple-negative subtypes with low *RUNX1* expression and that the increased *FOXO* activity supports continued cell proliferation and tumor progression [107, 115].

RUNX2 is located on chromosome 6p21 from positions 45,296,053 to 45,518,818, and produces instructions for transcription of a protein that builds and maintains the skeleton, specifically, osteoblast cells, suggesting that it is a regulator of 'bone genes'[116]. This gene is the least studied of the three family members, although it has been suggested that growth factor families involved in tumor cell growth, such as *FGF* and *IGF*, may signal through *RUNX2* and amplify its expression. *RUNX2* has been found to be amplified in osteosarcoma, however, more clinical studies need to be conducted to generalize this finding to other cancers [9]. The main interest in this gene is the implication that it is involved in bone metastasis. The potential role of *RUNX2* in bone metastasis was first observed when its target gene, *collagenase-3*, was expressed in breast cancer cell lines (MDA-MB-231), and whose cells are found to form osteolytic lesions in mice [117]. Breast cancer cells secrete parathyroid-hormone-related peptide (*PTHrP*), which encourages formation of osteoblasts during bone metastasis. *RUNX2* was reported to regulate *PTHrP* expression of metastatic breast cancer cells in the bone and the cell cycle of the cancer cells themselves [118-119]. It has also been demonstrated that *RUNX2* alters factors, which can facilitate metastasis, including *VEGF* [120]. One recent study evaluated the percentage of *RUNX2* immunoreactivity (positive protein expression)

within the nuclei of breast cancer cells and found it to be correlated with stage and histological grade of the tumor. It was also associated with recurrence and overall survival in patients with ER-negative tumors, but not ER+, suggesting a potential indicator of prognosis for specific subtypes although further evidence is required to assess these complex interactions [121].

RUNX3 is located on chromosome 1p36 from positions 25,226,002 to 25,291,612, approximately 66 kb long including 6 exons [122]. In normal cells *RUNX3* regulates cell proliferation and cell death by apoptosis by interacting with DNA repair proteins, inhibiting angiogenesis, and functioning in cell adhesion and invasion [123].

RUNX3 expression is known to be down-regulated in several cancers, including gastric, bile duct, pancreatic, colorectal, and lung, strengthening its role as a tumor suppressor in normal cells [124]. It has been observed that *RUNX3* is consistently under-expressed in breast cancer cells compared to normal breast epithelial cells and is most likely a result of protein mislocalization and hypermethylation of the promoter region of *RUNX3* [108, 125]. Fujii and colleagues reported silencing of *RUNX3* in breast cancer cells lines via a method of epigenetic mechanism described as: enhancer of Zeste Homologue 2 (EZH2)—mediated histone methylation of H3 at the Lys27 (H3K27) residue, which results in repressed transcription. In the same study, Fujii, et al. also found that EZH2 binds to the *RUNX3* promoter, resulting in increased H3K27 methylation and a subsequent decrease of *RUNX3* expression [126]. In contrast, *RUNX3* overexpression is found to be correlated with reduced metastasis of breast cancer cells [125]. One of the most common epigenetic pathway events in cancer, hypermethylation of the CpGisland, suggests that *RUNX3* inactivation is a significant risk factor for

carcinogenesis. In a group of bladder cancer patients, researchers found that *RUNX3* inactivation not only occurs early in the process but also increases with age [127].

To date, there does not appear to be any published population-based studies, which have evaluated SNPs on the *RUNX* genes and their association with breast cancer. The evidence to date for these genes is based on studies involving mouse models, gene expression, and copy number variants.

Crosstalk between TGF-β and ERα Signaling Pathways

It is widely acknowledged that the *ERα* and *TGF-β* signaling pathways help to regulate mammary development, function, and carcinogenesis. However, *TGF-β* family members and *ERα* have conflicting roles in cell proliferation and survival of breast epithelial cells. *ERα* signaling supports proliferation and differentiation specifically by enhancing cyclin D1 and c-Myc, components of the cell cycle. In contrast, *TGF-β* signaling pathway promotes apoptosis by reducing the expression of c-Myc and cyclin-dependent kinases in epithelial cells [128]. The opposing regulatory effect on cell proliferation has motivated researchers to evaluate the relationship between the two signaling pathways.

Several studies have provided evidence that *R-SMAD2, R-SMAD3* and *Co-SMAD4* come into direct physical contact with *ERα* [129-132]. *Co-SMAD4* is found to be a mediator of crosstalk between *TGF-β* and *ERα* where it acts as a co-repressor of the transcription of *ERα*, inhibiting tumor growth. Interestingly, *Co-SMAD4* has been found to induce apoptosis in ER+ but not ER– cells [133]. *R-SMAD3* has been found to be a co-activator of *ERα* changing the role of *TGF-β*. In the absence of *Co-SMAD4*, *TGF- β* can regulate *ERα* transcription through a *R-SMAD3*-mediated process and enhance the

estrogen-*ERα* cell proliferation [131]. In fact, Araki and colleagues found that 65% of late stage breast cancers are characterized by activated *R-SMAD3* and *HDM2*, a negative regulator of the tumor suppressor p53 [134]. Bierie and colleagues compared breast cancer expression signatures to *TGF-β* response signatures and found two correlations: first, the *TGF-β* response signature was associated with ER- tumors and poor prognosis; second, the absence of the *TGF-β* response signature was found to be higher in ER+ tumors and was associated with a poor prognosis [135].

On the other hand, Ito and colleagues reported that the ligand activated estrogen-*ERα* complex appears to cause degradation of *R-SMAD2/3* complex, thereby reducing *TGF-β* signaling. This degradation was not dependent on DNA binding or transcription of *ERα* and this non-genetic process was suggested to reduce migration and invasion caused by *TGF-β*; however, there are conflicting reports on the ability of *ERα* to effectively degrade *R-SMAD2/3* complexes [130]. *ERα* activation has been reported to inhibit *TGF-β* transcription activity by up to 60% in reporter assays [129-130]. The hypothesized mechanism includes estrogen acting directly on the *TGF-β* signaling pathway to block the phosphorylation of *R-SMAD 2/3* complex via ubiquitin-proteasome pathway [130].

The *RUNX* transcription factors are also involved in regulation of cell growth and differentiation and have been shown to bind with *SMADs*, which in turn can affect *ERα* transcription. *RUNX1* has been called an 'accessibility factor' for *ERα* binding sites and may function to establish a cooperative chromatin structure in DNA used for binding to EREs and to control gene expression in ER+ cells specifically.

Using mouse models, Huang and colleagues showed that *RUNX3* may target *ERα* to function as a tumor suppressor by destabilizing the gene and inhibiting the expression [136]. They also found an inverse relationship between the two genes' expression; the higher the *RUNX3* expression the lower the ERα in ER+ cells and *vice versa*, while Lau and colleagues found the same effect in ERα – cells [125, 136]. This supports the tumor suppressor role in *RUNX3* with *ERα* as a potential mediator of these effects.

Summary and Rationale for Research

It is evident that breast carcinogenesis is highly complex and many factors can vary by ethnicity. *ERα* has an imperative role in normal breast development as well tumor initiation and promotion. Its status is a relevant prognostic and predictive factor after breast cancer diagnosis. Genes in the *TGF-β* signaling pathway have numerous and often conflicting cellular effects, as either tumor promoters or suppressors, and as an either inhibitors or stimulators of angiogenesis. Through animal models, these genes and *ERα* have been found to have an effect on expression of one another through a *SMAD*-mediated process in the cell nucleus.

This evidence provides justification for studying genetic variation within these genes and the relationship between *TGF-β* and *ERα* signaling pathways as well as to evaluate whether the association differs in Hispanic and NHW women. While high penetrance mutations only account for 5% of breast cancers overall, it is believed that the combined effect of low-penetrance variants may explain a large component of breast cancer risk [137-138]. It is therefore important to continue to explore variation within and between genes and/or pathways as sources for the missing contribution to breast cancer susceptibility.

METHODS

Study Population and Data Collection

The study population consists of 11,060 participants in a multi-site consortium including data harmonized across three population-based case-control studies conducted within the U.S. and Mexico: the 4-Corner's Breast Cancer Study, the San Francisco Bay Area Breast Cancer Study, and the Mexico Breast Cancer Study. The goal of this consortium is to evaluate the biological basis of ethnic-related health disparities between Hispanic and NHW women for breast cancer risk and survival using genetic factors in the CHIEF signaling pathway in combination with behavioral, social, and cultural factors. It is hypothesized that these genetic factors may influence differences in breast cancer development and survival among these two ethnic groups [8]. A total of 40 selected SNPs, including those on *TGF-β1, TGF-βR1, RUNX1, RUNX2* and *RUNX3* were genotyped for the BCHD study. An additional five SNPs on *ERα* were genotyped for a subset of the participants from the New Mexico site (n=1,954) of the 4-CBCS as well for this dissertation. Taken together, these data were used to address the specific aims of this dissertation study. Details of methodologies specific to each study have been previously described [139-141]; brief descriptions including objectives, eligibility, recruitment, participation, and data collection are given below.

4-Corner's Breast Cancer Study

4-Corner's Breast Cancer Study (4-CBCS) was conducted from 1999-2005. The primary objective of this study was to evaluate the variation of risk factors associated with breast cancer between Hispanic and NHW women in the Southwest 4-Corners region of the U.S. Eligibility criteria for breast cancer cases included: self-identified Hispanic, American Indian, or NHW ethnicity; age 25 to 79 years; residency in Arizona (Cochise Coconino, Maricopa, Pima, Pinal, Santa Cruz, and Yuma counties), Colorado, New Mexico, or Utah; and diagnosis with a first primary breast cancer (*in-situ* or invasive) between October 1999 and May 2004. Cases were ascertained and eligibility was confirmed through the respective state cancer registries; New Mexico and Utah registries are a part of the SEER Program, and Arizona and Colorado registries are a part of the Centers for Disease and Control (CDC). At the time of selection, Hispanic ethnicity was identified using the computer program Generally Useful Ethnic Search System (GUESS) [142] and the Census Spanish Surname List [143]. NHW cases were matched on age to Hispanic cases and selected on a 1:1 ratio in Arizona and Colorado; a 4:1 ratio for all cases in Utah and for cases under age 50 years in New Mexico; and a 1:1 ratio for cases over age 50 years in New Mexico. Controls were frequency-matched based on 5-year age distributions and ethnicity of the cases and were randomly selected from the target populations using commercial mailing lists (Arizona and Colorado) and driver's license lists (New Mexico and Utah) for women under age 65, and from the Center for Medicare Services lists for women age 65 years and older. Potential participants self-reported their ethnicity in a telephone screening interview to determine eligibility.

A total of 5,163 women (cases: Hispanic=851, American Indian=22, NHW=1,683; controls: Hispanic=913, American Indian=23, NHW= 1,671) participated in the 4-CBCS. Trained interviewers administered a computer-assisted questionnaire in English or Spanish including questions regarding socio-demographics, medical and reproductive histories, family history of breast cancer, diet (modified to include foods common to Southwestern part of U.S.), physical activity, smoking, alcohol, medications,, and weight history. The referent year for most sections of the questionnaire was the year prior to diagnosis for cases and year prior to selection for controls. Anthropometric measures (weight, height, waist/hip circumference) were also recorded. For cases, the median time from diagnosis to date of interview was 671 days for Arizona; 540 days for Colorado; 599 days for New Mexico; and 267 days for Utah. Blood was also collected from the majority (75%) of participants (cases=1,244 NHW, 606 Hispanic; controls=1,329 NHW, 728 Hispanic) and DNA was extracted for subsequent analysis [139].

San Francisco Bay Area Breast Cancer Study

The San Francisco Breast Cancer Study (SFBCS) was conducted from 1995-2004. The primary objective of this study was to investigate the role of lifestyle factors, migration patterns and acculturation in Hispanic women, and ethnic differences in risk factors with breast cancer risk. A primary focus was on vitamin D exposure and SNPs in the vitamin D receptor gene. Eligibility criteria for breast cancer cases included: NHW, Hispanic, or African American ethnicity; age 35 to 79 years; living in the San Francisco Bay Area; and diagnosed with a first primary breast cancer (invasive) between April 1, 1995 and April 30, 1998 for phase 1, May 1, 1998 and April 30, 1999 for phase 2a, May 1, 1997 and April 30, 1998 for phase 2b, and May 1, 1999 and April 30, 2002 for phase

3. Cases were ascertained and eligibility was confirmed through the California State Cancer Registry, a SEER registry. Controls were randomly selected from the target population using random digit-dialing and frequency-matched to cases based on the expected 5-year age distributions and ethnicity. Potential participants self-reported ethnicity in a telephone screening interview to determine eligibility. A total of 3,823 women (cases= NHW=596, Hispanic=1,119; controls= NHW=646, Hispanic=1,462) participated in an in-person interview that was conducted with trained interviewers who administered a questionnaire in English or Spanish and took anthropometric measures (weight, height, waist/hip circumferences). Questions were similar to those used in the 4-CBCS with additional questions on occupational history and sunlight exposure. Blood specimens (or saliva sample, if blood collection was refused) were collected between 1999-2002 (phase3) for 1,105 (93%) cases an 1,318 (92%) of controls, and DNA was extracted for subsequent analysis [140].

Mexico Breast Cancer Study

The Mexico Breast Cancer Study (MBCS) was conducted from 2004-2007. The primary objective of this study was to investigate lifestyle, genetic, and socio-demographic factors associated with the risk of breast cancer in Mexican women. Eligibility criteria for breast cancer cases included: 35 to 69 years of age; living in Mexico City, Monterrey, or Veracruz metropolitan areas at least 5 years prior to selection; and diagnosed with a first primary breast cancer (*in-situ* or invasive) between January 2004 and December 2007. Cases were recruited from 12 participating hospitals in the three areas. Controls were selected using a probabilistic multistage design based on the basic geo-statistical catchment areas of the 12 participating hospitals and were frequency-matched to cases based on the expected 5-year age distributions, to a

healthcare institute membership, and residency. A total of 2,074 Mexican women (cases=1,000, controls=1,074) participated in an in-person interview that was conducted by trained interviewers. Anthropometric measures (weight, height, waist/hip circumferences) and mammograms were collected. For cases, the median time from diagnosis to date of interview was three days. Questions were similar to 4-CBCS with additional questions on number and type of possessions owned (i.e. telephones, stoves, televisions, computers, etc) that was used to construct a socioeconomic index based on distribution of these variables in controls using principle components analysis. Blood was collected for 850 (85%) cases and 1,031 (96%) controls and DNA was extracted for subsequent analysis [141].

All studies were reviewed and approved by their respective Institutional Review Board for Human Subjects and all participants signed a written informed consent prior to participation.

Harmonization of Data across Studies

Data from all three studies were harmonized at the University of Utah, the consortium's coordinating site, using questionnaire data from each study. Variables of interest based on study hypothesis involving the CHIEF pathway were identified and the distributions of data across studies were compared. The distribution of the variables was highly correlated ensuring the validity of the harmonization [8]. Table 2 summarizes each case-control study involved in the BCHD study.

Table 2: Descriptive summary of studies in the multi-collaborative case-control BCHD study *(n=11,060)*

Study Description	MBCS	SFBCS	4-CBCS
Study Objective	• to investigate lifestyle, genetic, and socio-demographic factors with the risk of breast cancer in Mexican women	• to investigate the role of lifestyle factors, migration patterns and acculturation in Hispanic women and ethnic differences in risk factors with breast cancer risk.	• to evaluate the variation of risk factors associated with breast cancer among Hispanic and NHW women living in the Southwestern part of the United States
Data collection period	• 2004-2007	• 1995-2004 (3 phases)	• 1999-2005
Case Eligibility Criteria and Ascertainment	• Hispanic ethnicity • resident of Mexico City, Monterrey, or Veracruz metropolitan areas at least 5 years prior to selection • age 35-69 y • diagnosed with a first primary breast cancer (*in-situ* or invasive) between January 2004 and December 2007 • recruited from 12 participating hospitals in the 3 areas	• NHW, Hispanic, or African American ethnicity • resident of the San Francisco Bay Area • age 35-79 y • diagnosed with a first primary breast cancer (invasive) between 1995-2002 • ascertained from SEER cancer registry	• Hispanic, American Indian or NHW ethnicity • resident of Arizona (Cochise Coconino, Maricopa, Pima, Pinal, Santa Cruz, and Yuma counties), Colorado, New Mexico, or Utah • age 25-79 y • diagnosed with a first primary breast cancer (in-situ or invasive) between October 1999 and May 2004 • ascertained from state cancer registries (SEER/CDC)
Control Eligibility Criteria and Ascertainment	• selected using a probabilistic multistage design based on the basic geo-statistic catchment area from the 12 participating hospitals • frequency-matched based on the expected 5-year age distributions, membership to a healthcare institute, and residency of the cases	• randomly selected using random digit-dialing • frequency-matched to cases based on the 5-year age distributions and ethnicity	• women <65 y were randomly selected using commercial mailing lists (Arizona and Colorado) and driver's license lists (New Mexico and Utah) • women age ≥65 y were selected from Center for Medicare Services list • frequency-matched to cases based on 5-year age distributions and ethnicity
Participant Sample Size	• cases=1,000 • controls=1,074 • total:2,074 Mexican women	• cases=1,715 (H=1,119, NHW=596) • controls=2,108 (H=1,462, NHW=646) • total: 3,823 women	• cases=2,556 (H=851, AI=22, NHW=1,683) • controls=2,607 (H=913, AI=23, NHW= 1,671) • total: 5,163 women
DNA Sample Size (case/control)	• 850/1,031 • 91% of total population	• 1,105/1,318 • 63% of total population • collected between 1999-2002 (phase3)	• 1,850/2,057 • 75% of total population

SNP Selection Criteria and Genotyping Methods

Selection of SNPs

With the exception of *ERα*, the genes of interest in this dissertation study were genotyped as part of the larger BCHD study. In the BCHD study, a tagSNP approach was utilized to define variation across candidate genes and SNPs were selected to be included in the platform using the following five parameters:

1. Linkage Disequilibrium (LD) blocks defined using a Caucasian LD map and an r^2 (pairwise LD measure representing correlation) ≥0.8 (when r^2=1, two SNPs are in perfect LD;
2. Minor allele frequency (MAF) >0.1;
3. Range of -1500 base pairs (bps) from the initiation codon to +1500 bps from the termination codon;
4. 1 SNP/LD bin; and
5. Functional status (i.e. determined via *in-vitro* studies)

ERα SNPs were selected based on a MAF>0.1 plus one or more of the following criteria:

1. Previous literature indicating an association with breast cancer or any other cancer
2. tagSNP
3. LD block ≥0.8 with SNP under investigation or with a known functional variant

In the current study, 45 SNPs among 6 genes were investigated:

1. *TGF-β1* (n=2): rs1800469, rs4803455;
2. *TGF-βR1* (n=5): rs6478974, rs1571590, rs1013186, rs11568785, rs10733710;
3. *RUNX1* (n=8): rs7279383, rs2268288, rs2252585, rs11701453, rs8127225, rs1474479, rs1883066, rs7279123;

4. *RUNX2* (n=17): rs1321075, rs17209895, rs2677108, rs2819854, rs2790093, rs9463090, rs2396441, rs1316330, rs7750470, rs6930053, rs12208240, rs12209785, rs10948238, rs13201287, rs12333172, rs1200428, rs598953;
5. *RUNX3* (n=8): rs2236850, rs9438876, rs7517302, rs906296, rs7551188, rs6688058, rs11249206, rs447876*;* and
6. *ERα* (n=5): rs2046210, rs6913578, rs851984, rs1801132, rs3798577.

Genotyping – BCHD Study

DNA was extracted from either whole blood (n=7,286) or saliva samples (n=637) from study participants. Whole Genome Amplification (WGA) was carried out on the saliva-derived DNA samples prior to genotyping. Genotyping included 1,466 SNPs in 205 candidate genes and was conducted in the coordinating site lab at the University of Utah using a multiplexed bead array assay format based on GoldenGate chemistry (Illumina, San Diego, California). A genotyping call rate of 99.93% was attained (99.65% for WGA samples). There were 132 internal replicates that were blinded, representing 1.6% of the sample set. The duplicate concordance rate was 99.996% as determined by 193,297 matching genotypes among sample pairs.

An additional 104 Ancestral Informative Markers (AIMs) were genotyped and used to characterize genetic admixture among participants [8]. AIMS were selected based on known differences in prevalence of alleles in Native American and European populations [144]. The computer program, STRUCTURE 2.0, was utilized to calculate the proportion of genetic admixture based on a two-population model that included European and Native American ancestry. STRUCTURE allows the multilocus genetic data to define the population structure. The proportion of an individual's genome

(ancestry) originating from one of two original founding populations is estimated using a clustering algorithm [145].

Genotyping – ERα for 4-CBCS New Mexico Site

Preparation for assays began with measuring the concentration (ng/μL) of each DNA sample using the Nanodrop2000 instrument and software (Thermo Scientific, Wilmington, DE, USA). The measured DNA concentration was used to determine the amount needed to dilute and standardize each sample to 4ng/μL for storage in the working plates.

ERα genotypes for 5 SNPs were determined using TaqMan assays (Applied Biosystems (ABI), Foster City, CA, USA) and evaluated on a 96-well single block ABI Step One Plus real-time polymerase chain reaction (PCR) machine. Each 20ul reaction sample contained 5uL genomic DNA (20ng), primers, probes, and TaqMan Universal Master Mix (containing AmpErase UNG, AmpliTaq Gold enzyme, dNTPs, and reaction buffer). Polymerase Chain Reaction (PCR) was carried out under the following conditions: 50°C for 2 minutes to activate UNG, 95°C for 10 min to active Gold enzyme, followed by 40 cycles of 92°C for 15 seconds and 60°C for 1 minute, and then through a final extension phase at 72°C for 7 minutes. A Veriti Thermal Cycler was used to run two 96-well plates simultaneously; however the ABI PCR machine reported the endpoint reading and produced results because of the specific feature called sequence detection, which gives fluorescent endpoint readings of the TaqMan-DNA reactions. Results were merged to the harmonized BCHD dataset via study ID for the New Mexico site analysis.

Description of Independent, Dependent Variables and Potential Confounders

The *TGF-β1, TGF-βR1, RUNX1, RUNX2, RUNX3,* and *ERα* SNPs were evaluated based on their genotypes. As an example, using alleles that make up the genotypes, C

represents the wild-type allele, and T represents the variant/minor allele so that each genotype was categorized: CC=0, CT=1, TT=2. Homozygous wild types were used as the referent group. Heterozygous and homozygous minor allele genotypes were analyzed for association of breast cancer risk. Women were excluded from analysis if missing data on any one particular SNP.

Diagnosis of breast cancer was the dependent or outcome variable. Cases were included in analysis if the baseline diagnosis was a first primary breast cancer, either *in-situ* (stage 0) or invasive (stages 1-4). Women were excluded from analysis if their baseline diagnosis was not a first primary breast cancer.

Potential confounders and the specified categories were selected based on *a priori* knowledge when there was evidence that a factor was a significant confounder of the main effect variable (i.e. SNPs) being modeled. Covariates that were assessed included:

- age (years, <40-referent, 40-49, 50-59, 60-69, 70+);
- study site (4-CBCS-referent, MBCS, SFBCS);
- self-reported ethnicity (NHW-referent, Hispanic/American Indian);
- first-degree family history of breast cancer (no-referent, yes);
- history of HRT (no-referent, yes);
- history of OC use (no-referent, yes);
- menopausal status (pre-/peri- vs. post-);
- age at menarche (years, <12-referent, 12, 13, 14+);
- parity (nulliparous-referent, 1-2, 3-4, 5+);
- age at FFTB (years, <20-referent, 20-24, 25-29, 30+);
- education (< high school-referent, high school/GED, post-high school);

- long-term alcohol consumption (grams per day, none-referent, low (<5), moderate (5-9), high (10+));
- physical activity (total hours of vigorous intensity activity/week, continuous);
- BMI during referent year (World Health Organization cut-points for normal-referent (18.5-24.9 kg/m^2), overweight (25-29.9 kg/m^2), and obese (30+ kg/m^2));
- cigarette smoking (never-referent, former or current status); and
- genetic admixture, % Native American ancestry (≤28%-referent, 0.29-0.70%, 71-100%)

BMI calculated from a participant's weight and height using the formula weight in kilograms divided by height in meters squared (kg/m^2). BMI is assessed as described by the well-known World Health Organization international cut-points for normal, overweight, and obese. Alcohol intake (grams per day) over the lifetime was available for all but 600 cases and controls from California. For these individuals, alcohol consumption during the referent year was used. Physical activity was measured as total hours of vigorous intensity per week. Menopausal status was determined based on responses to questions on menstrual history and hormone therapy use. Women who reported having periods during the referent year were defined as pre-menopausal. Women who reported using hormone therapy were defined as post-menopausal if they reported natural menopause (did not report have a period within the past 12 months) and were ≥95th percentile of age for race/ethnicity within each study site of those. This age for natural menopause was 58 years for NHW and 56 years for Hispanics from the 4-CBCS, 54 years for MBCS, and 55 years for NHW and 56 years for Hispanics from the SFBCS. Each covariate was assessed for its individual association with breast cancer risk, using a conservative significance level of (Wald p=≤0.20) to determine whether to further

evaluate it as a confounder in multivariable modeling. Participants were classified by level of percent Native American ancestry in genetic admixture groups based on evaluation of AIMs (See *Genotyping*). Cut-points were made based on the distribution of genetic ancestry in the control population so that each ancestral group had sufficient power when assessing associations.

Statistical Analysis

All statistical analyses were performed using SAS version 9.3 (SAS Institute, Cary, NC).

Descriptive Statistics

Descriptive statistics for all potential confounders, as well as the distribution of genotypes for selected SNPs for *RUNX1* (n=8), *RUNX2* (n=16), *RUNX3* (n=8), *TGF-β1* (n=2), *TGFβ-R1* (n=5) and *ERα* (n=5) genes, were calculated and reported by ethnicity and case-control status within each ethnic group. Significant differences between groups were determined using t-tests for continuous variables and chi-square (X^2) for categorical variables. Significant differences between study centers were also evaluated. Mantel Haenszel X^2 p-values for between and within group comparisons, were reported.

Testing for Hardy Weinberg Equilibrium

Genotype distributions were also evaluated for agreement with Hardy Weinberg Equilibrium (HWE) by the Pearson X^2 test among controls, which is used to compare the observed versus the expected frequency of genotypes [146]. HWE is an approximation of genotype frequencies in a population. The fundamental concept for HWE is that allele frequencies do not change from generation to generation, and assumes independent segregation of alleles at a locus. When HWE is satisfied, the following assumptions are made in a population: random mating, no mutations, and no selection, migration, or drift. The null hypothesis (H_0) is that HWE holds in a population, the alternate hypothesis (H_A)

is that it does not. In order to correct for multiple hypothesis testing, HWE was adjusted for the False Discovery Rate as originally described by Benjamini and Hochberg (1995) [147].

Assessment of Potential Confounders and Model Building

The best approach taken for statistical model building is one that minimizes confounding through evaluation of selected covariates and finds the most parsimonious model that best fits the data [148]. For the present analysis, a purposeful selection algorithm was utilized for multivariable model building as proposed originally by Hosmer and Lemeshow [50]. Each covariate was tested in a univariable logistic regression model; testing the individual association with risk of breast cancer. Genotypes for each SNP were initially assessed as a co-dominant (Let C=major allele and T=minor allele, (0 (CC) *vs.* 1(CT); and 0 (CC) *vs.*2 (TT)) mode of inheritance, adjusting for age and study site. Unconditional logistic regression modeling was used because the controls were frequency-matched, not fully matched, with the cases. OR and corresponding 95% confidence intervals (95% CI) are reported, along with Wald p-values. A covariate with a Wald p-value ≤0.2 was considered a candidate for inclusion when constructing the multivariable models. The multivariable logistic regression model was based on the following equation:

$$logit\,\{P(Y=1)\} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_k X_k \qquad (1)$$

Let $P(Y=1)$= *log* (probability of being a case over probability of being a control); X=covariate(s) of interest.

When constructing multivariable models, full models were compared to a reduced model (age and study site adjusted OR for the particular SNP) to evaluate the presence of confounding.

Those corresponding to a >10% change in the effect estimate (OR) for the SNP, at a significance level of α=0.05, were retained in the final multivariable logistic regression models after further calculating the log likelihood ratio of models with and without the significant covariates [50]. The following is an example of a full model with all potential covariates of interest:

(2)

$$logit\,\{P(Y=1)\} = \beta 0 + \beta 1(ER\alpha\ (rs2046210, (CC\ vs\ CT,\ \ TT))) + \beta 2(age) + \beta 3(study\ center) + \beta 4\ (ethnicity) + \beta 5(menopausal\ status) + \beta 6\ (BMI) + \beta 7(genetic\ admixture) + \beta 8(HRT) + \beta 9(parity) + \beta 10(physical\ activity) + \beta 11(alcohol) + \beta 12(smoking) + \beta 13(education) + \beta 14(OCU) + \beta 15(age\ at\ menarche) + \beta 16(age\ at\ menopause) + \beta 17(family\ history) + \beta 18(age\ at\ FFTB)$$

Statistical models of the associations for the SNPs were also evaluated based on mode of inheritance as follows: (Let C=major allele and T=minor allele) additive model (0 (CC) + 1 (CT) + 2 (TT), continuous); dominant model (0 (CC) *vs.* 1 (CT) + 2 (TT) (referent)); or recessive model (0 (CC) + 1(CT) (referent) *vs.* 2 (TT)), when sufficient power could be gained by collapsing genotypes and a trend towards a different mode of inheritance is observed for OR for significant models using a co-dominant mode of inheritance.

Effect Modifiers and Interaction

Multiplicative interaction models were used to determine the statistical interactions between the 45 SNPs of interest with the hypothesized effect modifiers including menopausal status and proportion of Native American ancestry.

Using menopausal status as an example, the interaction model can be described as:

(3)

$$logit\ \{P(Y=1)\} = \beta_0 + \beta_1(SNP) + \beta_2(menopausal\ status) + \beta_2(SNP * menopausal\ status) + \cdots + \beta_k X_k$$

The significance of the interaction term was tested using the X^2 p-value for interaction ($p < 0.05$). The interaction term was further evaluated using the difference in maximum likelihood estimates for logistic regression models, with and without the constructed interaction term, using a X^2 test with 2 degree of freedom (2-df, co-dominant model) or 1 degree of freedom (1-df, additive, dominant, or recessive models). Evaluation of odds ratios between the strata was conducted and p values for heterogeneity were calculated for differences in association between strata groups by comparing the difference in maximum likelihood estimates for a logistic regression model as described above. Trend p-values were conducted based on the X^2 p-value of the SNP assessed as a continuous variable within each stratum. SNP-SNP interactions, including those with a potentially meaningful association with breast cancer ($p<0.15$), were conducted and evaluated in the same manner. A multiplicative interaction effect on the logit scale was assumed for these two-way interactions.

Multinomial Logistic Regression

Utilizing data from cases and controls, multinomial logistic regression was conducted to evaluate the association between genotypes and risk of developing breast tumor phenotypes defined by ER/PR status (+/-). This statistical technique is an extension of logistic regression where the dependent variable has >2 categories, also known as a polytomous response. For the BCHD study this method was based on a nominal dependent variable with five unordered categories: control (referent); cases

(ER+/PR+, ER+/PR-, ER-/PR+, and ER-/PR-). For the New Mexico subpopulation this method was based on a nominal dependent variable with three unordered categories: control (referent); case with ER+ tumor; and case with ER- tumor. Using the maximum likelihood estimation, the probability of being a case with an ER+ or ER- tumor was compared to the probability of being a control, creating several binary logistic regression models [50, 149]. Additionally, the hypothesized effect modifiers, menopausal status and genetic admixture, were assessed via stratification, although sample size and power was reduced. A p-value for heterogeneity was calculated using an interaction term for each model. Trend p-values were conducted based on the X^2 p-value of the SNP assessed as a continuous variable within each stratum. Equation 4 depicts the separate binary logistic regression models that results from using the multinomial model fit.

$$logit \quad \{P(y = ER+)/P(y = control)\} = \beta_0 + \beta_1(SNP) + \beta_2 x_2 \ +\cdots+ \beta_K x_K \qquad (4)$$

$$logit \quad \{P(Y = ER-)/P(y = control)\} = \beta_0 + \beta_1(SNP) + \beta_2 x_2 \ +\cdots+ \beta_K x_K$$

Genetic Risk Score-Cumulative effect of risk alleles

While these SNPs may have a relatively weak individual effect on breast carcinogenesis, there may be a stronger cumulative effect. The hypothesize for the present study is that these SNPs work with one another for proper function of the gene and therefore the cumulative effect may have an association with breast cancer.

SNPs that were significantly ($p<0.05$) or marginally significantly ($p<0.15$) associated with breast cancer (per-allele/trend effect) were used to create a genetic risk score (GRS). These SNPs were then divided into two groups: those that increased or decreased risk. The two GRS were calculated by treating each risk allele equally,

simply counting the number of risk alleles for each SNP (0, 1, or 2) and summing across SNPs [150]. GRS was created for the *TGF-β* signaling pathway genes (*RUNX1, RUNX2, RUNX3, TGF-β1, TGF-βRI*); and across *TGF-β* signaling pathway genes and *ERα.* When not evaluated as a continuous variable, GRS categorization was based on the distribution of the variable among controls and differs for each of the two GRS. Multivariable logistic regression analyses were conducted and OR and 95% CI, adjusting for age, study, and genetic admixture, were estimated for each GRS category. Trend p-values were also conducted evaluating GRS as a continuous variable.

Multiple Comparisons

The p-values, for both main effects and interactions, based on 1-df Wald X^2 test statistic, were adjusted for multiple comparisons taking into account tagSNPs within each gene using the step-down Bonferroni-correction method [151]. This method is based on the effective number of independent SNPs as determined using the SNP spectral decomposition method proposed by Nyholt [152], and modified by Li and Ji [153] using the eigenvalues of a correlation matrix of the SNPs for each gene. This method of correcting for multiple comparisons is conservative, especially when evaluating correlated variables such as SNPs within a gene. However, it is less conservative than the conventional Bonferroni correction because you have more opportunities to reject the null hypotheses, which results in an increase in statistical power [153].

Power Analysis

Calculating the power of a case-control genetic association study is necessary to determine if there is a sufficient sample size to detect a hypothesized effect, or know the smallest detectable effect, between groups. Power is the probability of successfully detecting an effect, or difference, between groups. It is calculated from the false-

negative, or type II, error rate (Beta, B) as 1-B. Eighty percent is the most common sought after power. This means that there is an 80% or greater chance of finding a statistically significant difference, if one is present, or of rejecting the null hypothesis, when it should be, thus avoiding a type II error rate [146]. Statistical power depends upon: sample size and the specified magnitude of effect (i.e. odds ratio/relative risk), genotype frequencies, and desired level of statistical significance (alpha, α=0.05 (the false positive, type I, error rate)). There are several assumptions made when calculating power, including disease prevalence in the general population compared to population under investigation, penetrance of alleles (corresponding to mode of inheritance), high linkage disequilibrium (LD or D´) with disease loci (D´=0.8), and proportion of variance explained by the loci under investigation [146, 154].

Power was calculated utilizing software called *The Genetic Power Calculator* (available at: http://pngu.mgh.harvard.edu/~purcell/gpc/cc2.html) for the analysis of discreet traits in case-control studies; in the present study, the presence or absence of breast cancer [154]. The following parameters were fixed based on the studies' participation and blood collection: number of cases, and ratio of controls to cases. The estimate for prevalence of breast cancer was based on the SEER age-adjusted prevalence of breast cancer for all races [2]. Table 3 shows the variables used for known or fixed parameters used in all power analyses for the entire study population. The software was utilized to estimate power assuming varying frequencies for MAF (determined from genotyping), and genotype relative risk (assuming an additive model: risk increases 2-fold when there are 2 minor alleles) for the entire study population, as well as for the New Mexico site. The power is reported for a 1-df (A *vs.* a) and a general 2-df test (AA *vs.* Aa *vs.* aa).

Table 3: Fixed case-control parameters for the BCHD study and the New Mexico site of the 4-CBCS.

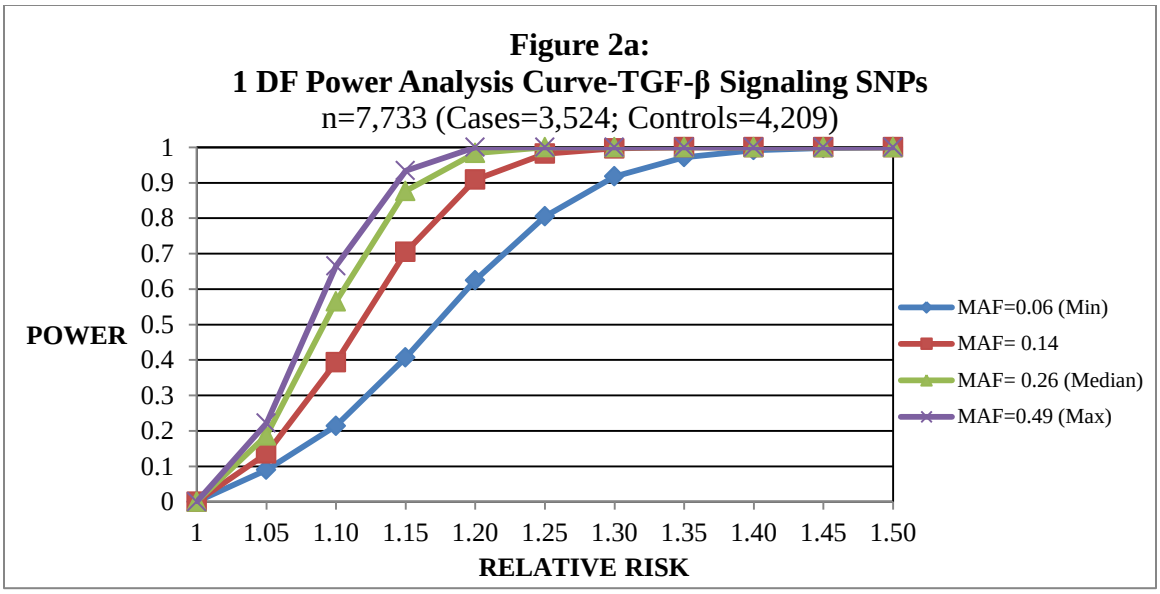
	BCHD Study Population			
Parameter	Total	Hispanic	non-Hispanic White	New Mexico [a]
Number of cases	3,524	2,093	1,431	694
Control : case ratio	1.194	1.247	1.118	1.03
SEER Prevalence of breast cancer	0.0108	0.0108	0.0108	0.0108
D prime (D´)	0.80	0.80	0.80	0.80
Defined Type I error rate (α)	0.05	0.05	0.05	0.05
Defined Power	0.80	0.80	0.80	0.80

[a] Number of cases for New Mexico are a subset of the number of cases for the BCHD study population.

Because 40 selected SNPs were evaluated from *TGF-β1, TGF-βR1, RUNX1, RUNX2,* and *RUNX3* genes in this population, power was estimated using the minimum, median, and maximum MAF of the SNPs taken together. In addition, a 'threshold' for a specified relative risk was determined for varying MAF and ethnic-specific subgroups. Power analysis was conducted for the three studies combined and stratified by ethnicity (See Table 3). There were five SNPs from *ERα* and power was evaluated based on the 4-CBCS New Mexico site participants with available DNA samples (n=1,458).

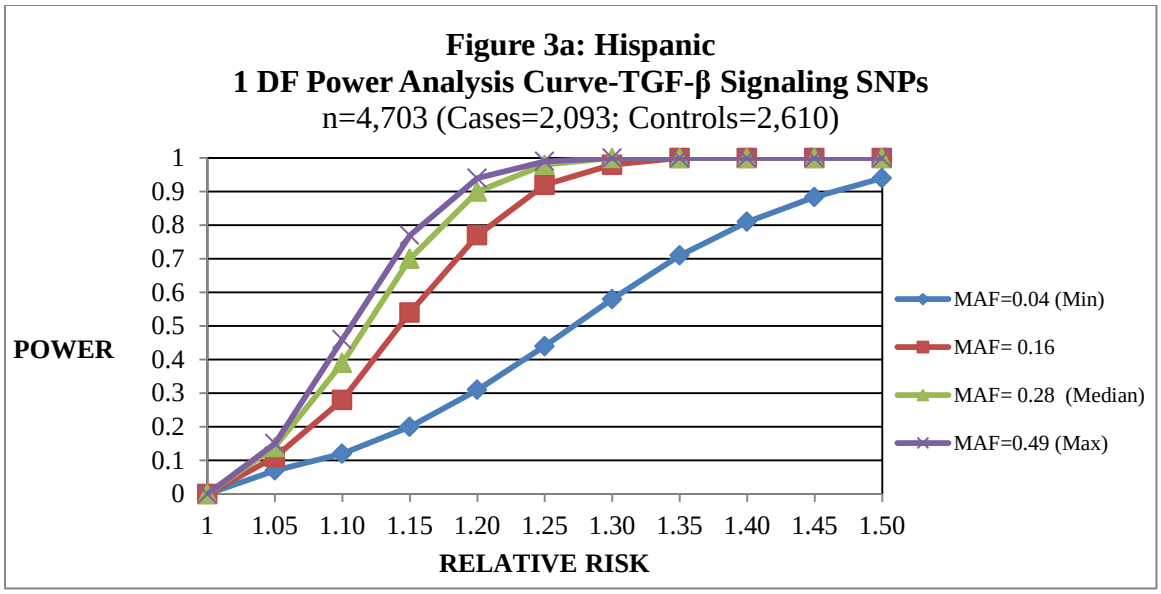
Results of Power Analysis

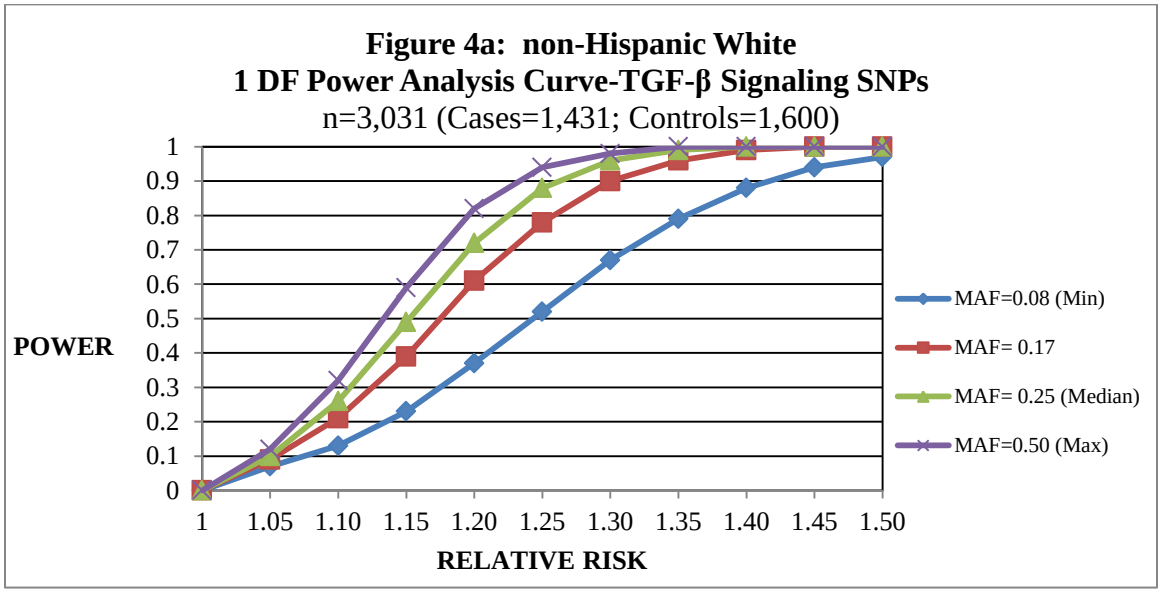
Figure 2a shows the power curve for the whole study population based upon varying MAF of TGF-β signaling SNPs of interest when RR=1.05-1.50, measured in 0.5 increments. Assuming a 1-df model, there was 80% power to detect the following: RR=1.10 when MAF=0.49; R=1.12 when MAF=0.26; RR=1.17 when MAF=1.14; and RR=1.25 when MAF=0.06.



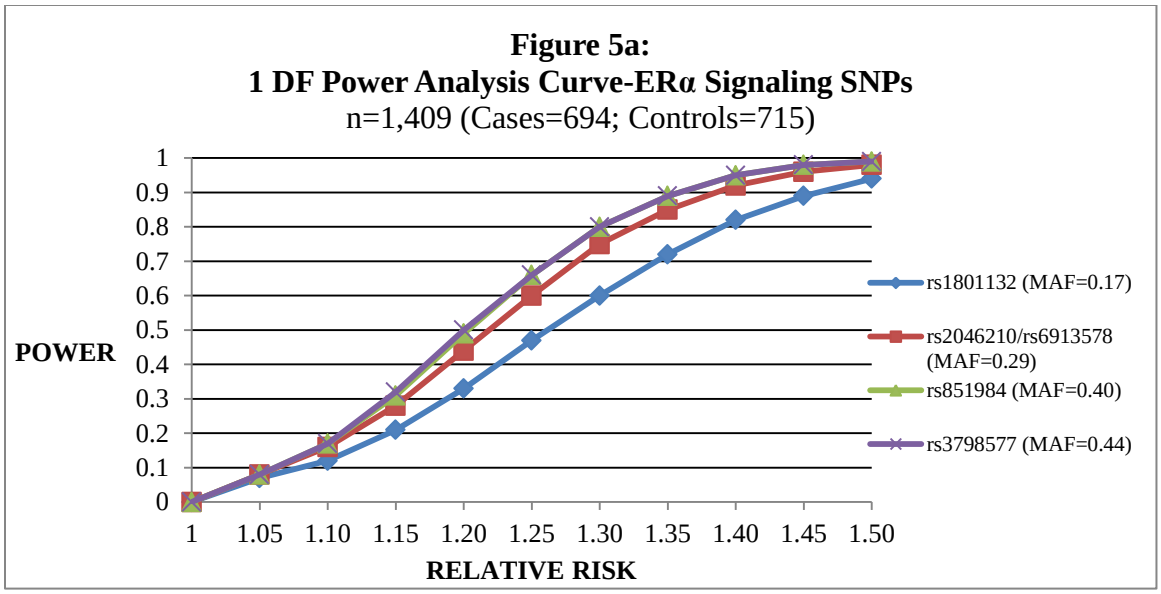
[a] DF=degree of freedom; MAF= Minor allele frequency.

MAF for each SNP can differ by subgroups, which is evident for a few SNPs when stratified by ethnicity. As an example, both Hispanic and NHW women had the lowest MAF for *TGF-βR1* rs11568785; however NHW was higher than for Hispanic women: 8% *versus* 4%, respectively. Figure 3a shows the power curve for Hispanic women. For 80% power, assuming a 1-df model, the present study was able to detect a RR=1.15 when MAF=0.49; R=1.17 when MAF=0.28; RR=1.20 when MAF=0.16; and RR=1.40 when MAF=0.04. The NHW subgroup has a smaller sample size; therefore the detectable RRs are slightly higher than for the Hispanic subgroup for similar MAF. The present study was able to detect a RR=1.19 when MAF=0.50; R=1.22 when MAF=0.25; RR=1.25 when MAF=0.17; and RR=1.35 when MAF=0.08 (Figure 4a). Results are similar when evaluating power using a 2-df model (Figures 2b, 3b, 4b in *Appendix*).

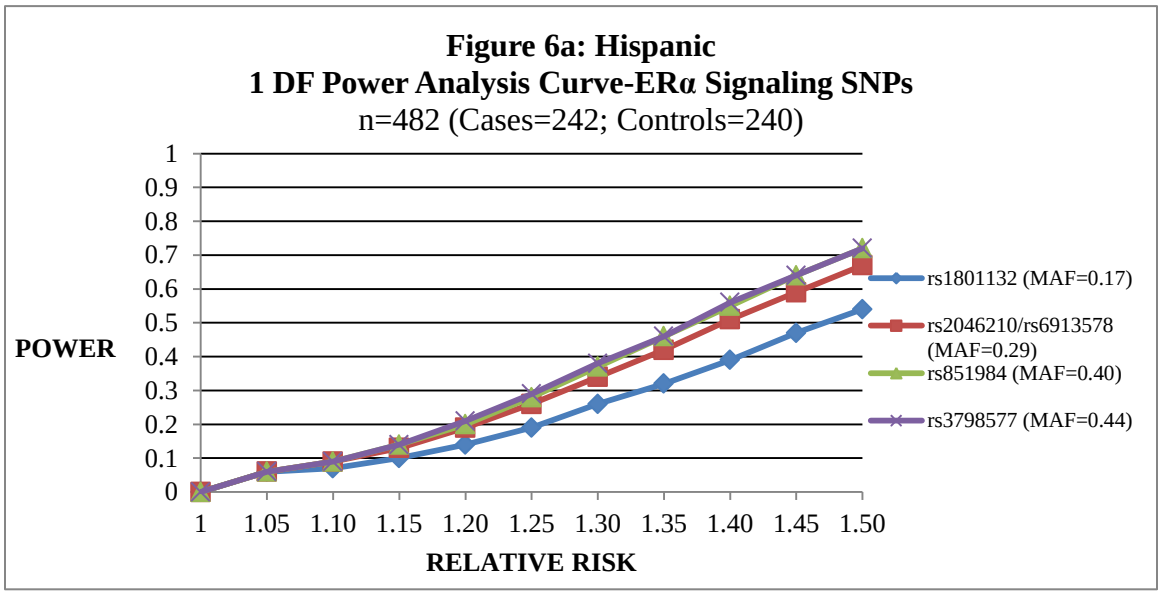




The power curve for the overall New Mexico sample using MAF for *ERα* SNPs is shown in Figure 5a. For 80% power, assuming a 1df model, the present study had the ability to detect a RR=1.30 when MAF=0.44 or 0.40; RR=1.33 when MAF=0.29; and RR=1.40 when MAF=0.17.



There were approximately twice as many NHW women compared to Hispanic women in this New Mexico subsample (927 *vs.* 482). For Hispanic women, regardless of MAF, there was not sufficient power (80%) to detect a RR of 1.05-1.50; with only 70% power to detect RR=1.50 (Figure 6a). In contrast, for NHW women there was 80% to detect a RR=1.38 when MAF=0.44 or 0.40; RR=1.40 when MAF=0.29; and RR=1.50 when MAF=0.17 (Figure 7a). Results are similar when evaluating power using a 2-df model in this subsample (Figures 5b, 6b, 7b in *Appendix*).



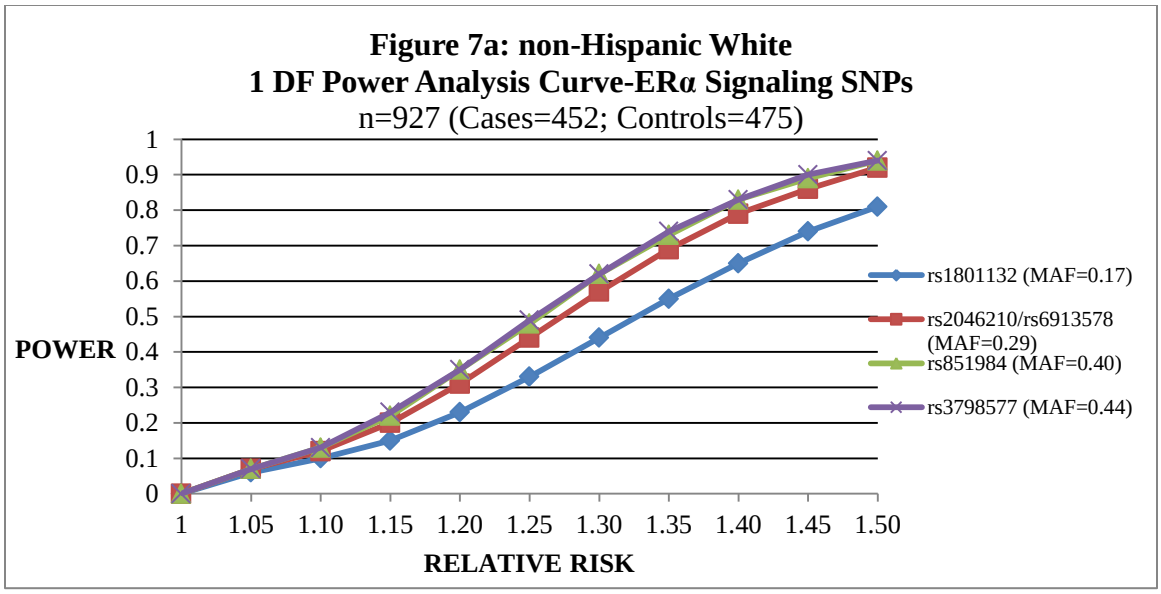


Figure 7a: non-Hispanic White
1 DF Power Analysis Curve-ERα Signaling SNPs
n=927 (Cases=452; Controls=475)
POWER
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
1
1.05
1.10
1.15
1.20
1.25
1.30
1.35
1.40
1.45
1.50
RELATIVE RISK
rs1801132 (MAF=0.17)
rs2046210/rs6913578 (MAF=0.29)
rs851984 (MAF=0.40)
rs3798577 (MAF=0.44)

RESULTS

Descriptive Characteristics - BCHD

The present study included a total of 7,733 participants (cases=3,524; NHW=1,431, Hispanic=2,093; controls=4,209; NHW=1,599, Hispanic=2,610) (Table 4) that had available DNA (69.9%). Statistically significant differences were observed by ethnic group regardless of case-control status for all descriptive variables, with the exception of age at menarche (*p-values* in Table 4). Approximately 60% of all participants were over 50 years of age; with a higher proportion of pre-menopausal women among Hispanic than NHW women for both cases (41% *vs.* 34%) and controls (41% *vs.* 32%) ($p<0.0001$).

Among NHW women, >99% of participants were characterized by the lowest NA ancestry strata (≤0.28), while 65% of Hispanic women had a moderate (0.28-0.70) proportion of NA ancestry. Among, Hispanic compared to NHW women tended to be more obese (≥30 kg/m²) (43% *vs.* 26%), be less educated (<high school) (57% *vs.* 5%), report higher parity (5+) (21% *vs.* 9%), and to report no history of smoking (71% *vs.* 58%) (data not shown, *p-values* in Table 4). Hispanic cases were more frequently diagnosed with ER- tumors compared to NHW cases (26% *vs.* 20%) (Table 4). NHW compared to Hispanic cases reported a higher alcohol intake (≥10 g/day) (13% *vs.* 3%), HRT (71% *vs.* 57%), and have more family history of breast cancer (23% *vs.* 12%) (Table 4).

Among statistically significant differences within Hispanic women, cases compared to controls had a higher proportion of 1st degree family history of breast cancer (12% *vs.* 8%*)*; age at menarche, ≤12 years of age (49% *vs.* 45%); and history of HRT (32% *vs.* 29%). Conversely, Hispanic compared to NHW cases had a lower proportion of participants with number of full-term births (parity 3+) (52% *vs.* 63%); age at first full-term birth <20 years of age (24% *vs.* 32%); obesity, BMI 30or greater kg/m^2 (39.7% *vs.* 45.5%); and less than a high school education (53% *vs.* 60%). Hispanic controls were more likely to have >70% ancestry than cases (24% *vs*. 21%) (Table 4). Among statistically significant differences within NHW women, cases compared to controls tended to have a first degree family history of breast cancer (23% *vs.* 16%) and report OCU (71% *vs.* 66%). NHW controls tended to be >70 years of age (19% *vs.* 14%); and reported more full-term births (3+) (44% *vs.* 38%) compared to NHW cases (Table 4).

These potential confounders were further tested for differences by study sites. The 4-CBCS contributed 82.9% of NHW and 27.9% of Hispanic women in the BCDS, while the SBCS contributed 17.1% of NHW and 33.5% of Hispanic women. The MBCS contributed 39% of Hispanic women. The goal of this additional testing was to evaluate to what extent the differences in characteristics between case-control status were influenced by study site. There were no differences between study sites for age at first full-term birth, OCU and physical activity. However, statistically significant differences between study sites were observed for other covariates under evaluation. For example, there was a statistically significant greater proportion of cases compared with controls who reported a parity ≥5 within the studies: MBCS (20% and 29%); SFBCS (13% and 20%); 4-CBCS (10% and 13%) as well as across site ($p<0.0001$) (data not shown).

Table 4. Descriptive covariates stratified by ethnicity and case-control status, *Breast Cancer Health Disparities Study (n=7,733)*

	non-Hispanic White (n= 3,030)					Hispanic/American Indian (n= 4,703)					
	Cases (n=1431)		Controls (n= 1599)			Cases (n= 2093)		Controls (n=2610)			
Covariate, categorical	*N*	(%)	*N*	(%)	p[a]	*N*	(%)	*N*	(%)	p[a]	p[b]
Age (years)											
<40	87	6.1	117	7.3		198	9.5	313	12.0		
40-49	401	28.0	409	25.6		708	33.8	834	31.9		
50-59	403	28.2	410	25.6	0.04	614	29.3	758	29.0	0.22	<.0001
60-69	340	23.8	356	22.3		425	20.3	530	20.3		
70+	200	13.9	307	19.2		148	7.1	175	6.7		
Study											
4-CBCS	1177	82.3	1335	83.5		579	27.7	736	28.2		
MBCS	-	-	-	-	0.37	816	39.0	994	38.1	0.94	<.0001
SFBCS	254	17.8	264	16.5		698	33.4	880	33.7		
Family history, 1st degree											
Yes	312	22.6	237	15.5	<.0001	244	11.9	208	8.2	<.0001	<.0001
No	1071	77.4	1289	84.5		1799	88.1	2326	91.8		
Menopausal status											
Pre-/ peri-menopausal	475	34.1	494	31.5	0.13	831	41.2	1027	40.7	0.71	<.0001
Post-menopausal	919	65.9	1075	68.5		1186	58.8	1499	59.3		
Age at menarche (years)											
<12	282	20.1	288	18.3		496	24.0	555	21.6		
12	389	27.7	435	27.6	0.12	523	25.3	608	23.6	0.002	0.99
13	381	27.1	418	26.5		466	22.6	590	22.9		
14+	353	25.1	435	27.6		581	28.1	820	31.9		
Parity											
Nulliparous	249	17.6	248	15.7		225	10.8	181	7.0		
1-2	622	44.0	638	40.3	0.0005	774	37.2	790	30.5	<.0001	<.0001
3-4	441	31.2	529	33.4		729	35.0	997	38.4		

	non-Hispanic White (n= 3,030)					Hispanic/American Indian (n= 4,703)					
	Cases (n=1431)		Controls (n= 1599)			Cases (n= 2093)		Controls (n=2610)			
Covariate, categorical	*N*	(%)	*N*	(%)	p[a]	*N*	(%)	*N*	(%)	p[a]	p[b]
≥5	102	7.2	167	10.6		353	17.0	626	24.1		
Age at first full-term birth (years)											
Nulliparous	249	17.6	248	15.7		225	10.9	181	7.0		
<20	170	12.0	199	12.6		503	24.4	825	31.9		
20-24	499	35.3	609	38.5	0.91	716	34.7	922	35.7	0.005	.0005
25-29	314	22.2	342	21.6		356	17.3	451	17.5		
30+	180	12.8	184	11.6		262	12.7	204	7.9		
Body mass index (kilograms/meter^2)											
<25	651	46.1	699	44.4	0.30	482	23.4	453	17.6		
25-29.9	411	29.1	465	29.5		762	37.0	951	36.9	<.0001	<.0001
30+	350	24.8	412	26.1		818	39.7	1172	45.5		
History of hormone replacement therapy											
Ever	795	69.2	899	69.2	0.98	614	32.4	709	29.3	0.03	<.0001
Never	354	30.8	401	30.8		1282	67.6	1708	70.7		
History of oral contraceptive use											
Ever	996	71.1	1035	65.8	.0018	1174	56.8	1413	54.8	0.16	<.0001
Never	405	28.9	539	34.2		892	43.2	1167	45.2		
Alcohol intake (g/day)											
None	573	49.6	691	53.3		1531	84.0	1926	85.8		
Low (<5g/day)	292	25.3	315	24.3	0.09	148	8.1	175	7.8	0.18	<.0001
Moderate (5-<10g/day)	135	11.7	130	10.0		96	5.3	80	3.6		
High (≥10g/day)	155	13.4	161	12.4		47	2.6	65	2.9		
Smoking status											
Never	649	56.0	794	60.3		1280	70.3	1616	72.1		
Former	368	31.7	360	27.3	0.13	313	17.2	347	15.5	0.38	<.0001
Current	143	12.3	163	12.4		228	12.5	278	12.4		

	non-Hispanic White (n= 3,030)					Hispanic/American Indian (n= 4,703)					
	Cases (n=1431)		Controls (n= 1599)			Cases (n= 2093)		Controls (n=2610)			
Covariate, categorical	*N*	(%)	*N*	(%)	p[a]	*N*	(%)	*N*	(%)	p[a]	p[b]
Education											
<High school	71	5.0	79	5.0		1088	52.8	1538	60.2		
High school grad/GED	284	20.1	338	21.3	0.57	377	18.3	419	16.4	<.0001	<.0001
Post High school	1059	74.9	1168	73.7		594	28.9	597	23.4		
% Native American ancestry											
≤ 0.28	1420	99.2	1591	99.5		276	13.2	280	10.7		
0.28-0.70	7	0.5	7	0.44	0.21	1373	65.6	1697	65.0	.0012	<.0001
>0.70	4	0.3	1	0.1		444	21.2	633	24.3		
Estrogen/Progesterone Receptor (ER/PR) Status											
ER+/PR+	681	68.4	-	-		598	61.8	-	-		
ER+/PR-	116	11.7	-	-	-	117	12.1	-	-	-	0.001
ER-/PR+	17	1.7	-	-		28	2.9	-	-		
ER-/PR-	181	18.2	-	-		225	23.2	-	-		
Covariate, continuous	Mean	SD	Mean	SD	p[c]	Mean	SD	Mean	SD	p[c]	p[d]
Age	55.6	11.2	56.7	12.3	0.05	52.7	10.6	52.3	10.8	0.26	<.0001
Study, Total MET hrs/week											
4-CBCS	1.35	2.9	1.44	3.1	0.47	1.18	2.9	1.00	2.4	0.22	.0008
MBCS	-	-	-	-	-	1.45	4.1	1.48	5.9	0.89	-
SFBCS	1.94	3.10	2.21	4.0	0.39	1.09	2.7	0.97	3.1	0.44	<.0001

Note: Column percentages (%) may not add up to 100% due to rounding. Column totals (n) may not add up to total for each column due to missing observations: education (n=121), family history (n=247), menopausal status (n=227), age at menarche (n=113)parity (n=62), age at first full-term birth (n-=94), HRT use (n=971), OCU (n=112), alcohol (n=1,213), smoking status (n=1,194), BMI (n=107), vigorous physical activity (n=56).

[a] Case-control status comparisons within each ethnic group, *Mantel-Haenszel* chi-square p values reported

[b] Ethnic group comparisons regardless of case-control status, *Mantel-Haenszel* chi-square p values reported

[c] Case-control status comparisons within each ethnic group (and studies for physical activity), p values from t- tests reported

[d] Ethnic group comparisons regardless of case-control status, p values from t- tests reported

Descriptive Characteristics for 4-CBCS – New Mexico Site

A total of 927 NHW (cases=452, controls=475) and 482 Hispanic women (cases=242, controls=240) with available DNA (72.1%) were included in analyses evaluating SNPs from the *ERα* gene, based on data from the New Mexico site only. Descriptive characteristics of the New Mexico subset of the BCHD study population stratified by ethnicity and case-control status are shown in Table 5. Similar to the larger study population, a larger proportion (63%) of participants were over 50 years of age; with a slightly higher proportion of post-menopausal women for NHW compared to Hispanic (67% *vs.* 63%) women. In addition to being slightly older, NHW women, compared to Hispanic women, tended to report more family history of breast cancer (22% *vs.* 16%), lower parity (nulliparous) (18% *vs.* 10%), less obesity (≥30 kg/m²) (21% *vs.* 31%), history of HRT (69% *vs.* 57%), history of OCU (69% *vs.* 58%), no long-term alcohol consumption (46% *vs.* 64%), history of cigarette smoking (48% *vs.* 38%), post-high school education (76% *vs.* 50%), ≤0.28 NA ancestry (99% *vs.* 20%), and a higher mean vigorous MET hours per week of physical activity (1.45 *vs.* 1.09) (data not shown, *p-values* in Table 5). There were only 13 participants total that had >70% NA ancestry; therefore the strata were collapsed into two categories (≤28%, 29-70%). The presence of a family history of breast cancer was significantly higher for both NHW cases (27% *vs.* 18%) and Hispanic cases (21% *vs.* 10%) compared to controls (Table 5). There was no difference for ER/PR status by ethnicity (p=0.13), which was observed in the larger population (Table 5).

Table 5. Descriptive covariates stratified by ethnicity and case-control status,
New Mexico sub-population of 4-Corners Breast Cancer Study (n=1,409)

	non-Hispanic White (n=927)					Hispanic/American Indian (n=482)					
	Cases (n=452)		Controls (n=475)			Cases (n=242)		Controls (n=240)			
Covariate, categorical	*N*	(%)	*N*	(%)	p[a]	*N*	(%)	*N*	(%)	p[a]	p[b]
Age (years)											
<40	19	4.2	20	4.2		21	8.7	23	9.6		
40-49	134	29.7	138	29.1		85	35.1	79	32.9		
50-59	109	24.1	129	27.2	0.86	67	27.7	66	27.5	0.92	<.0001
60-69	121	26.8	113	23.8		50	20.7	60	25.0		
70+	69	15.3	75	15.8		19	7.9	12	5.0		
Family history, 1st degree											
Yes	115	26.5	81	17.8	0.002	48	21.4	22	9.8	.0007	0.005
No	319	73.5	373	82.2		176	78.6	203	90.2		
Menopausal status											
Pre-/ peri-menopausal	151	33.4	151	31.8	0.60	91	37.6	85	35.6	0.64	0.13
Post-menopausal	301	66.6	324	68.2		151	62.4	154	64.4		
Age at menarche (years)											
<12	88	19.6	77	16.3		54	22.3	52	21.8		
12	106	23.7	115	24.4	0.17	68	28.1	58	24.3	0.66	0.10
13	139	31.0	141	29.9		52	21.5	63	26.4		
14+	115	25.7	139	29.5		68	28.1	66	27.6		
Parity											
Nulliparous	89	19.7	74	15.6		22	9.1	27	11.3		
1-2	223	49.3	215	45.4	0.006	99	40.9	93	38.8	0.75	<.0001
3-4	121	26.8	154	32.5		90	37.2	90	37.5		
≥5	19	4.2	31	6.5		31	12.8	30	12.5		

	non-Hispanic White (n=927)					Hispanic/American Indian (n=482)					
	Cases (n=452)		Controls (n=475)			Cases (n=242)		Controls (n=240)			
Covariate, categorical	*N*	(%)	*N*	(%)	p[a]	*N*	(%)	*N*	(%)	p[a]	p[b]
Age at first full-term birth (years)											
Nulliparous	89	19.7	74	15.6		22	9.1	27	11.3		
<20	54	12.0	66	13.9		50	20.7	60	25.0		
20-24	150	33.2	188	39.7	0.97	106	43.8	98	40.8	0.08	0.18
25-29	98	21.7	92	19.4		42	17.4	44	18.3		
30+	61	13.5	54	11.4		22	9.1	11	4.6		
Body mass index (kilograms/meter^2)											
<25	214	47.7	229	48.4		82	34.2	73	30.7		
25-29.9	143	31.9	142	30.0	0.95	87	36.3	89	37.4	0.42	<.0001
30+	92	20.5	102	21.6		71	29.6	76	31.9		
History of hormone replacement therapy											
Ever	248	69.5	263	69.2	0.94	96	53.0	113	61.8	0.09	<.0001
Never	109	30.5	117	30.8		85	47.0	70	38.2		
History of oral contraceptive use											
Ever	310	68.7	325	68.9	0.97	137	56.9	139	58.2	0.77	<.0001
Never	141	31.3	147	31.1		104	43.1	100	41.8		
Alcohol intake (g/day)											
None	195	43.1	233	49.3		159	66.3	145	61.4		
Low (<5g/day)	128	28.3	123	26.0	0.10	42	17.5	44	18.6	0.19	<.0001
Moderate (5-<10g/day)	66	14.6	58	12.3		18	7.5	18	7.6		
High (≥10g/day)	63	13.9	59	12.5		21	8.8	29	12.3		
Smoking status											
Never	226	50.0	257	54.3		154	63.9	143	59.8		
Former	151	33.4	144	30.4	0.24	58	24.1	57	23.9	0.21	0.006
Current	75	16.6	72	15.2		29	12.0	39	16.3		

	non-Hispanic White (n=927)					Hispanic/American Indian (n=482)					
	Cases (n=452)		Controls (n=475)			Cases (n=242)		Controls (n=240)			
Covariate, categorical	*N*	(%)	*N*	(%)	p[a]	*N*	(%)	*N*	(%)	p[a]	p[b]
Education											
<High school	21	4.7	17	3.6		54	22.4	39	16.2		
High school grad/GED	83	18.3	105	22.2	0.63	68	28.2	81	33.8	0.34	<.0001
Post High school	348	77.0	352	74.3		119	49.4	120	50.0		
% Native American ancestry											
≤ 0.28	441	98.9	463	99.6		50	21.1	43	17.9		
0.28-0.70	3	0.6	2	0.4	0.15	181	76.4	192	80.0	0.49	<.0001
>0.70	2	0.5	0	0.0		6	2.5	5	2.1		
Estrogen/Progesterone Receptor (ER/PR) Status											
ER+/PR+	200	68.7	-	-		91	61.1	-	-		
ER+/PR-	33	11.3	-	-	-	18	12.1	-	-	-	0.13
ER-/PR+	3	1.0	-	-		6	4.0	-	-		
ER-/PR-	55	18.9	-	-		34	22.8	-	-		
Covariate, continuous	Mean	SD	Mean	SD	p[c]	Mean	SD	Mean	SD	p[c]	p[d]
Age	56.5	11.1	56.5	11.2	0.97	53.1	10.7	52.8	10.7	0.81	<.0001
Total MET hrs/week	1.40	3.3	1.51	3.4	0.61	0.90	1.89	1.30	3.3	0.11	0.03

Note: Column percentages (%) may not add up to 100% due to rounding. Column totals (n) may not add up to total for each column due to missing observations: education (n=2), family history (n=72), menopausal status (n=1), age at menarche (n=8), parity (n=1), age at first full-term birth (n=1), HRT lifetime (n=308), OCU lifetime (n=6), alcohol (n=8), smoking status (n=4), BMI (n=9), genetic admixture (n=21), or ER/PR status (n=254; cases only included, n=440).

[a] Case-control status comparisons within each ethnic group, *Mantel-Haenszel* chi-square p values reported

[b] Ethnic group comparisons regardless of case-control status, *Mantel-Haenszel* chi-square p values reported

[c] Case-control status comparisons within each ethnic group (and studies for physical activity), p values from t- tests reported

[d] Ethnic group comparisons regardless of case-control status, p values from t- tests reported

Descriptive data for SNPs in the *TGF-β* signaling pathway and *ERα*

A description of selected SNPs from *TGF-β* signaling pathway and *ERα* including: the SNP(s) relative to the chromosome (region, location, and position), major and minor alleles, MAF by ethnicity, HWE by ethnicity and proportion missing are shown in Table 6. The MAF and HWE are calculated based on the frequencies of alleles and genotypes in the control population. In the BCHD population, the MAF for the majority of SNPs was ≥ 0.10 in both NHW and Hispanic populations. NHW women had two SNPs with a MAF=0.08 (*RUNX2* (rs12208240) and *TGF-βR1* (rs11568785)). Hispanic women had one SNP with a MAF=0.04 (*TGF-βR1* (rs11568785)) and one SNP with a MAF=0.07 (*RUNX1* (rs1883066)). The HWE assumption ($p > 0.05$) was met for all SNPs analyzed in this dissertation. The proportion missing for each SNP is virtually zero, but it is important to note that the dataset was restricted to those with data for each SNP so that the sample size may differ for each analysis but differs by no more than 4 participants. The genotype distributions (homozygous wild-type, heterozygote, and homozygous variant) of the 45 SNPs evaluated are shown in Table 7 stratified by ethnicity and case-control status (*See Appendix*). Statistically significant differences were observed between ethnic groups for SNPs (3 genotypes), regardless of case-control status, for all but four SNPs (*RUNX2* (rs10948238 and rs7750470); *RUNX3* (rs4478762 and rs6688058)). For the majority of SNPs, the proportions of homozygous variants were higher among NHW compared to Hispanic women, with the exception of 11 SNPs: *RUNX1* (rs2252585 and rs8127225); *RUNX2* (rs1200428, rs12208240, rs12209785, rs13201287, rs1321075, rs2677108, rs598953); *TGF-β1* (rs1800469); and *TGF-βR1* (rs10733710), where Hispanic had a higher proportion of homozygous variants.

Table 6. Chromosome, allele, MAF and HWE for selected SNPs for genes in TGF-β signaling pathway, *Breast Cancer Health Disparities Study (n=7,733)* and ERα, *4-CBCS – New Mexico Site (n=1,409)*.

SNP	Chromosome			Alleles		MAF[1]		HWE[2]		Proportion Missing
	Region	Location	Position	Major	Minor	NHW	H/NA	NHW	H/NA	
RUNX1										
rs7279383	INTRON	21q22.3	36224963	C	G	0.19	0.11	0.89	0.74	0.0002
rs2268288	INTRON	21q22.3	36232671	T	C	0.20	0.12	1.00	0.84	0.0000
rs2252585	INTRON	21q22.3	36241929	T	C	0.27	0.43	0.96	0.42	0.0000
rs11701453	INTERGENIC	21q22.3	36338916	G	C	0.20	0.15	0.96	0.81	0.0000
rs8127225	INTERGENIC	21q22.3	36364765	T	C	0.13	0.27	0.96	0.75	0.0014
rs1474479	INTERGENIC	21q22.3	36405666	G	A	0.38	0.17	0.78	0.40	0.0000
rs1883066	INTERGENIC	21q22.3	36412156	G	C	0.12	0.07	0.96	0.91	0.0002
rs7279123	INTERGENIC	21q22.3	36415087	C	T	0.26	0.18	0.96	0.69	0.0033
RUNX2										
rs17209895	INTRON	6p21	45402445	T	C	0.27	0.14	0.96	0.08	0.0000
rs2677108	INTRON	6p21	45403774	T	C	0.41	0.54	0.97	0.72	0.0005
rs2819854	INTRON	6p21	45404528	T	C	0.51	0.47	0.96	0.43	0.0005
rs2790093	INTRON	6p21	45437484	A	G	0.33	0.31	0.86	0.68	0.0000
rs9463090	INTRON	6p21	45453345	G	A	0.21	0.18	0.86	0.82	0.0010
rs2396441	INTRON	6p21	45467765	C	T	0.50	0.49	0.86	0.11	0.0002
rs1316330	INTRON	6p21	45469626	G	T	0.25	0.16	0.66	0.72	0.0005
rs7750470	INTRON	6p21	45473256	T	C	0.19	0.20	0.97	0.84	0.0000
rs6930053	INTRON	6p21	45488758	C	T	0.41	0.30	0.89	0.99	0.0000
rs12208240	INTRON	6p21	45501937	G	A	0.08	0.12	0.96	0.95	0.0000
rs12209785	INTRON	6p21	45506122	A	G	0.25	0.28	0.59	0.52	0.0005
rs10948238	INTRON	6p21	45511541	C	T	0.39	0.39	0.68	0.32	0.0007
rs13201287	INTRON	6p21	45511945	G	A	0.25	0.30	0.62	0.64	0.0000
rs12333172	INTRON	6p21	45512215	C	T	0.20	0.16	1.00	0.52	0.0000
rs1200428	UTR	6p21	45518202	C	A	0.22	0.28	0.62	0.95	0.0000

SNP	Chromosome			Alleles		MAF[1]		HWE[2]		Proportion
	Region	Location	Position	Major	Minor	NHW	H/NA	NHW	H/NA	Missing
rs598953	INTERGENIC	6p21	45520030	T	A	0.37	0.43	0.93	0.71	0.0000
RUNX3										
rs2236850	INTRON	1p36	25240341	T	C	0.44	0.40	0.96	0.59	0.0017
rs9438876	INTRON	1p36	25241116	A	G	0.54	0.42	0.62	0.80	0.0000
rs7517302	INTRON	1p36	25254317	T	C	0.43	0.37	0.96	0.88	0.0007
rs906296	INTRON	1p36	25264658	C	G	0.23	0.18	0.96	0.39	0.0007
rs7551188	INTRON	1p36	25273200	C	T	0.54	0.47	0.98	0.21	0.0012
rs6688058	INTRON	1p36	25274998	G	A	0.13	0.14	0.96	0.61	0.0000
rs11249206	INTRON	1p36	25277982	T	C	0.51	0.35	0.96	0.81	0.0159
rs4478762	INTRON	1p36	25281015	G	A	0.11	0.12	0.62	0.82	0.0005
TGF-β1										
rs1800469	INTERGENIC	19q13.1	41860296	C	T	0.32	0.46	0.62	0.82	0.0124
rs4803455	INTRON	19q13.1	41851509	C	A	0.49	0.35	0.96	0.63	0.0589
TGF-βRI										
rs6478974	INTRON	9q22	101874403	T	A	0.47	0.35	0.89	0.94	0.0002
rs1571590	INTRON	9q22	101883808	A	G	0.20	0.10	0.97	0.47	0.0002
rs1013186	INTRON	9q22	101884337	G	A	0.20	0.11	0.97	0.60	0.0000
rs11568785	INTRON	9q22	101905834	A	G	0.08	0.04	0.96	0.67	0.0000
rs10733710	INTRON	9q22	101907424	G	A	0.23	0.35	0.97	0.12	0.0002
ERα										
rs2046210	INTERGENIC	6q25.1	151948366	G	A	0.36	0.28	0.85	0.85	0.02
rs6913578	INTERGENIC	6q24	151949556	A	C	0.33	0.25	0.95	0.85	0.01
rs851984	5'-UTR	6q25.1	152023191	G	A	0.38	0.39	0.95	0.95	0.007
rs1801132	EXON 4	6q24	152265522	C	G	0.22	0.26	0.85	0.95	0.006
rs3798577	3'-UTR	6q24	152462823	T	C	0.45	0.40	0.85	0.95	0.003

[1]Minor Allele Frequency (MAF) is based on control population

[2] Hardy-Weinberg Equilibrium (HWE) is based on the control population and is FDR adjusted

Overall, regardless of ethnicity, genotype distributions differed between cases and controls for *RUNX3* (rs906296, p=0.005, homozygous wild-type, CC, 64% *vs.* 61%); and *TGF-βR1* (rs10733710, p=0.01, homozygous variant, AA, .8.5% *vs.* 10%) and (rs6478974, p=0.01, homozygous variant, AA, 18% *vs.* 16%)).

Among NHW women, genotype distributions differed between cases and controls for only 2 SNPs: *RUNX2* (rs10948238, p=0.04, homozygous variant, TT, 18% *vs.* 14%) and *TGF-βR1* (rs10733710, p=0.03, homozygous variant, AA, 3.9% *vs.* 5%). Between cases and controls in Hispanic women there was a difference in genotype distributions for two SNPs: *RUNX2* (rs6930053, p=0.05, homozygous variant, TT, 11% *vs.* 8.9%) and *RUNX3* (rs906296, p=0.007, homozygous wild-type, CC, 63% *vs.* 67%).

In the New Mexico sub-population, genotype distributions differed by ethnicity in all but one of the *ERα* SNPs (rs851984, p=0.53), as shown in Table 8 (*See Appendix*). There were a higher proportion of homozygous variants among NHW cases and controls, for *ERα* SNPs: rs3798577, rs2046210, and rs6913578. Hispanic cases and controls had a higher proportion of homozygous variants for *ERα* (rs1801132). Genotype distributions did not differ between cases and controls within NHW or Hispanic women for *ERα* SNPs.

Univariable Analysis

Univariable OR(s), 95% CI(s), and p-values are reported for all descriptive characteristics for the BCHD study population and were further stratified by self-reported ethnicity in Table 9. Significant covariates, defined with p-values ≤ 0.20, associated with breast cancer in the overall study population were considered in multivariable modeling and included: age, family history, age at menarche, parity, age at first full-term birth,

BMI, history of HRT and OCU, long-term alcohol consumption, smoking status, education, and genetic admixture. Study site was also included in multivariable modeling as the distributions of significant covariates have been found to vary in each study population. Univariable OR(s) were comparable for majority of covariates stratified by Hispanic and NHW ethnicities, except for age (70+), [OR_H=1.34; 95% CI 1.01-1.77 *vs.* OR_{NHW}= 0.88; 95%CI: 0.63-1.22]; BMI (30+ kg/m^2), [OR_H=0.66; 95% CI 0.56-0.77 *vs.* OR_{NHW}= 0.91; 95%CI: 0.76-1.09]; history of HRT, [OR_H=1.15; 95% CI 1.01-1.31 *vs.* OR_{NHW}= 1.00; 95%CI: 0.84-1.19]; and education (< high school) [OR_H=0.71; 95% CI 0.62-0.82 *vs.* OR_{NHW}= 0.99; 95%CI: 0.71-1.38]. Table 10 reports the Univariable OR(s), 95% CI(s), and p-values for descriptive characteristics for the New Mexico sub-population and further stratified by self-reported ethnicity. Overall, only family history and education showed significant association with breast cancer risk, however, the following covariates are implicated as risk factors in previous literature and were considered in multivariable modeling: age, age at menarche, parity, age at first full-term birth, BMI, history of HRT and OCU, alcohol, smoking status, and genetic admixture. Associations stratified by Hispanic and NHW in this sub-population proved to be more divergent compared to the overall BCHD population, in particular for age (70+), [OR_H=1.73; 95% CI 0.68-4.41 *vs.* OR_{NHW}= 0.97; 95%CI: 0.48-1.97]; parity (5+) [OR_H=1.27; 95% CI 0.60-2.69 *vs.* OR_{NHW}= 0.51; 95%CI: 0.27-0.98]; age at first full-term birth (30+) [OR_H=2.45; 95% CI 0.98-6.14 *vs.* OR_{NHW}= 0.94; 95%CI: 0.58-1.52]; history of HRT [OR_H=0.70; 95% CI 0.46-1.06 *vs.* OR_{NHW}= 1.01; 95%CI: 0.74-1.39]; alcohol consumption (high ≥10 g/day) [OR_H=0.66; 95% CI 0.36-1.21 *vs.* OR_{NHW}= 1.28; 95%CI: 0.85-1.91]; and smoking status (current) [OR_H=0.69; 95% CI 0.41-1.18 *vs.* OR_{NHW}= 1.18;

95%CI: 0.82-1.71]. Genetic admixture was not associated with breast cancer in this subpopulation (Table 10).

Table 9. Univariable Odds Ratios (OR) and 95% Confidence Intervals (CI) for Descriptive Characteristics: *The Breast Cancer Health Disparities Study by Self-reported Ethnicity*

	Total			NHW (n=3,030)		Hispanic (n=4,703)	
Covariate, categorical	OR	95% CI	p	OR	95% CI	OR	95% CI
Age (years)							
<40	1.00	REF		1.00	REF	1.00	REF
40-49	1.35	1.14-1.59	.00006	1.32	0.97-1.80	1.34	1.09-1.65
50-59	1.31	1.11-1.56	0.002	1.32	0.97-1.80	1.28	1.04-1.58
60-69	1.30	1.09-1.56	0.004	1.28	0.94-1.76	1.27	1.02-1.58
70+	1.09	0.89-1.34	0.41	0.88	0.63-1.22	1.34	1.01-1.77
Study							
4-CBCS	1.00	REF		1.00	REF	1.00	REF
MBCS	0.97	0.87-1.08	0.57	--	--	1.04	0.91-1.20
SFBCS	0.98	0.88-1.09	0.73	1.09	0.90-1.32	1.01	0.87-1.17
Family history, 1st degree							
No	1.00	REF		1.00	REF	1.00	REF
Yes	1.57	1.38-1.80	<.0001	1.58	1.31-1.91	1.52	1.25-1.84
Menopausal status							
Pre-/ peri-menopausal	1.00	REF		1.00	REF	1.00	REF
Post-menopausal	0.95	0.87-1.05	0.31	0.89	0.76-1.04	0.98	0.87-1.10
Age at menarche (years)							
<12	1.00	REF		1.00	REF	1.00	REF
12	0.95	0.83-1.08	0.42	0.91	0.74-1.13	0.96	0.81-1.14
13	0.91	0.80-1.04	0.17	0.93	0.75-1.15	0.88	0.74-1.05
14+	0.81	0.71-0.92	0.001	0.83	0.67-1.03	0.79	0.68-0.93
Parity							
Nulliparous	1.00	REF		1.00	REF	1.00	REF
1-2	0.89	0.76-1.03	0.11	0.97	0.79-1.20	0.79	0.63-0.98
3-4	0.69	0.60-0.81	<.0001	0.83	0.67-1.03	0.59	0.47-0.73
5+	0.52	0.44-0.62	<.0001	0.61	0.45-0.82	0.45	0.36-0.57
Age at first full-term birth (years)							
Nulliparous	1.00	REF		1.00	REF	1.00	REF
<20	0.60	0.51-0.70	<.0001	0.85	0.65-1.11	0.49	0.39-0.61
20-24	0.72	0.62-0.84	<.0001	0.82	0.66-1.01	0.63	0.50-0.78
25-29	0.77	0.65-0.90	0.002	0.91	0.72-1.16	0.64	0.50-0.81
30+	1.03	0.85-1.25	0.75	0.97	0.74-1.28	1.03	0.79-1.35
Body mass index (kilograms/meter^2)							
<25	1.00	REF		1.00	REF	1.00	REF
25-29.9	0.84	0.75-0.94	0.003	0.95	0.80-1.13	0.75	0.64-0.88
30+	0.75	0.67-0.84	<.0001	0.91	0.76-1.09	0.66	0.56-0.77

	Total			NHW (n=3,030)		Hispanic (n=4,703)	
Covariate, categorical	OR	95% CI	p	OR	95% CI	OR	95% CI
History of hormone replacement therapy							
No	1.00	REF		1.00	REF	1.00	REF
Yes	1.13	1.03-1.24	0.01	1.00	0.84-1.19	1.15	1.01-1.31
History of oral contraceptive use							
No	1.00	REF		1.00	REF	1.00	REF
Yes	1.17	1.06-1.28	0.001	1.28	1.10-1.50	1.09	0.97-1.22
Alcohol intake (g/day)							
None	1.00	REF		1.00	REF	1.00	REF
Low (<5g/day)	1.12	0.97-1.29	0.12	1.12	0.92-1.36	1.06	0.85-1.34
Moderate (5-<10g/day)	1.37	1.13-1.66	0.002	1.25	0.96-1.63	1.51	1.11-2.05
High (≥10g/day)	1.11	0.91-1.36	0.30	1.16	0.91-1.49	0.91	0.62-1.33
Smoking status							
Never	1.00	REF		1.00	REF	1.00	REF
Former	1.20	1.07-1.36	0.003	1.25	1.05-1.50	1.14	0.96-1.35
Current	1.05	0.90-1.22	0.52	1.07	0.84-1.38	1.04	0.86-1.25
Education							
<High school	0.77	0.69-0.85	<.0001	0.99	0.71-1.38	0.71	0.62-0.82
High school grad/GED	0.93	0.82-1.06	0.27	0.93	0.78-1.11	0.90	0.76-1.08
Post high school	1.00	REF		1.00	REF	1.00	REF
% Native American ancestry							
≤ 0.28	1.00	REF		-	-	-	-
0.28-0.70	0.89	0.81-0.98	0.02	-	-	-	-
>0.70	0.78	0.68-0.90	.0004	-	-	-	-
Continuous Variable							
Total MET hrs/week	1.00	0.99-1.01	0.87	0.99	0.97-1.01	1.01	0.99-1.02

Missing observations: education (n=121), family history (n=247), menopausal status (n=227), age at menarche (n=113)parity (n=62), age at first full-term birth (n-=94), HRT use (n=971), OCU (n=112), alcohol (n=1,213), smoking status (n=1,194), BMI (n=107), vigorous physical activity (n=56).

Table 10. Univariable Odds Ratios (OR) and 95% Confidence Intervals (CI) for Descriptive Characteristics: ***New Mexico Sub-population by Self-reported Ethnicity***

	Total (n=1,409)			NHW (n=927)		Hispanic (n=482)	
Covariate, categorical	OR	95% CI	p	OR	95% CI	OR	95% CI
Age (years)							
<40	1.00	REF		1.00	REF	1.00	REF
40-49	1.09	0.68-1.74	0.73	1.02	0.52-2.00	1.18	0.61-2.29
50-59	0.97	0.60-1.56	0.90	0.89	0.45-1.75	1.11	0.56-2.20
60-69	1.06	0.66-1.72	0.80	1.13	0.57-2.22	0.91	0.45-1.84
70+	1.09	0.65-1.83	0.75	0.97	0.48-1.97	1.73	0.68-4.41
Family history, 1st degree							
No	1.00	REF		1.00	REF	1.00	REF
Yes	1.84	1.40-2.42	<.0001	1.66	1.20-2.29	2.51	1.46-4.33
Menopausal status							
Pre-/ peri-menopausal	1.00	REF		1.00	REF	1.00	REF
Post-menopausal	0.92	0.74-1.15	0.47	0.93	0.71-1.22	0.92	0.63-1.33
Age at menarche (years)							
<12	1.00	REF		1.00	REF	1.00	REF
12	0.91	0.67-1.26	0.58	0.81	0.54-1.21	1.13	0.67-1.89
13	0.85	0.62-1.16	0.31	0.86	0.59-1.27	0.80	0.47-1.35
14+	0.81	0.59-1.11	0.19	0.72	0.49-1.07	0.99	0.59-1.65
Parity							
Nulliparous	1.00	REF		1.00	REF	1.00	REF
1-2	0.95	0.70-1.30	0.75	0.86	0.60-1.24	1.31	0.70-2.45
3-4	0.79	0.57-1.09	0.15	0.65	0.44-0.96	1.23	0.65-2.31
5+	0.75	0.47-1.18	0.21	0.51	0.27-0.98	1.27	0.60-2.69
Age at first full-term birth (years)							
Nulliparous	1.00	REF		1.00	REF	1.00	REF
<20	0.75	0.52-1.09	0.13	0.68	0.42-1.09	1.02	0.52-2.01
20-24	0.81	0.59-1.12	0.21	0.66	0.46-0.97	1.33	0.71-2.48
25-29	0.94	0.66-1.34	0.72	0.89	0.58-1.35	1.17	0.58-2.37
30+	1.16	0.76-1.77	0.49	0.94	0.58-1.52	2.45	0.98-6.14
Body mass index (kilograms/meter^2)							
<25	1.00	REF		1.00	REF	1.00	REF
25-29.9	1.02	0.80-1.30	0.90	1.08	0.80-1.45	0.87	0.57-1.34
30+	0.93	0.72-1.22	0.62	0.97	0.69-1.35	0.83	0.53-1.31
History of hormone replacement therapy							
No	1.00	REF		1.00	REF	1.00	REF
Yes	0.88	0.69-1.13	0.32	1.01	0.74-1.39	0.70	0.46-1.06
History of oral contraceptive use							
No	1.00	REF		1.00	REF	1.00	REF
Yes	0.97	0.78-1.21	0.79	0.99	0.75-1.31	0.95	0.66-1.36
Alcohol intake(g/day)							
None	1.00	REF		1.00	REF	1.00	REF

	Total (n=1,409)			NHW (n=927)		Hispanic (n=482)	
Covariate, categorical	**OR**	**95% CI**	**p**	**OR**	**95% CI**	**OR**	**95% CI**
Low (<5g/day)	1.09	0.84-1.41	0.53	1.24	0.91-1.70	0.87	0.54-1.41
Moderate (5-<10g/day)	1.18	0.84-1.66	0.34	1.36	0.91-2.03	0.91	0.46-1.82
High (≥10g/day)	1.02	0.73-1.42	0.91	1.28	0.85-1.91	0.66	0.36-1.21
Smoking status							
Never	1.00	REF		1.00	REF	1.00	REF
Former	1.10	0.86-1.39	0.46	1.19	0.89-1.59	0.95	0.61-1.45
Current	0.99	0.73-1.33	0.93	1.18	0.82-1.71	0.69	0.41-1.18
Education							
<High school	1.35	0.94-1.96	0.03	1.25	0.65-2.41	1.40	0.86-2.26
High school grad/GED	0.82	0.64-1.05	0.02	0.80	0.58-1.11	0.85	0.56-1.28
Post high school	1.00	REF		1.00	REF	1.00	REF
% Native American ancestry							
≤ 0.28	1.00	REF		-	-	-	-
0.29-0.70	0.98	0.77-1.24	0.85	-	-	-	-
Continuous Variable							
Total MET hrs/week	0.98	0.94-1.01	0.20	0.99	0.95-1.03	0.93	0.87-1.02

Missing observations: education (n=2), family history (n=72), menopausal status (n=1), age at menarche (n=8), parity (n=1), age at first full-term birth (n=1), HRT lifetime (n=308), OCU lifetime (n=6), alcohol (n=8), smoking status (n=4), BMI (n=9), genetic admixture (n=21)

Multivariable Analyses

The TGF-β signaling (n=40) and ERα (n=5) SNPs were initially assessed as codominant models (data not shown). After evaluation of univariate OR (s), 95% CI (s) and p-values, the following were assessed as recessive models: RUNX1 (rs2252585, rs2268288, rs1474479); RUNX2 (rs12333172, rs12209785, rs10948238, rs13201287); RUNX3 (rs4478762, rs6688058, rs7517302); TGF-βR1 (rs6478974, rs10733710, rs11568785) and ERα (rs1801132, rs3798577); or as dominant models: RUNX1 (rs7279383, rs8127225, rs1883066); RUNX3 (rs906296); TGF-β1 (rs4803455) and ERα (rs2046210), while the remaining SNPs were kept in codominant models of inheritance. Six SNPs [RUNX2 (10948238), RUNX3 (rs906296), TGF-βR1 (rs6478974, rs10733710), and ERα (rs1801132, rs3798577)] were found to be independently associated with breast cancer risk ($p \leq 0.05$, data not shown).

When model building, a covariate was considered significant in multivariable analyses if it altered the age and study adjusted OR by ≥10%. Due to missing observations for each covariate multivariate model was restricted to women who had all data for all covariates so there was comparability for the same sample number. Confounding was not observed at this level so the characteristics [family history, age at menarche, parity, age at first full-term birth, BMI, history of HRT and OCU, long-term alcohol consumption, smoking status, or education] were not retained in the final models as covariates in analyses presented. Age and study (BCHD multi-site study only) were included as the base model and in subsequent analyses because characteristics related to the risk breast cancer differ among the two. Genetic admixture was also included as a covariate to account for differences in allele/genotype distribution between strata of NA ancestry.

Overall Association with Breast Cancer

After adjustment for age, study (TGF-β signaling SNPs only), and genetic admixture, nine SNPs [*RUNX1* (rs7279383 and rs8127225); *RUNX2* (rs10948238 and rs13201287); *RUNX3* (rs906296); *TGF-β1* (rs4803455); *TGF-βR1* (rs6478974), and *ERα* (rs1801132 and rs3798577)] were found to be significantly associated with overall breast cancer risk (Table 11a). Results for all SNPs are found in Table 11b (*See Appendix*).

A significant increase risk of breast cancer was observed with the dominant models of *RUNX3* (rs906296, CG/GG *vs.* CC, OR=1.15; 95% CI 1.04-1.26; p_{adj}=0.03) and *RUNX1* (rs8127225, TC/CC *vs.* TT, OR=1.11; 95% CI 1.01-1.22; p_{adj}=0.23) and recessive models of *RUNX2* (rs10948238, TT *vs.* TT/TC, OR=1.15; 95% CI 1.01-1.30; p_{adj}=0.42) and *ERα* [(rs1801132, GG *vs.* CC/CG, OR=1.72; 95% CI 1.10-2.69; p_{adj}=0.08)

and (rs3798577, TT *vs.* TT/TC, OR=1.36; 95% CI 1.04-1.76; p_{adj}=0.08)]. There was an inverse association with dominant model of *RUNX1* (rs7279383, CG/GG *vs.* CC, OR=0.89; 95% CI 0.80-0.99; p_{adj}=0.23) and *TGF-β1* (rs4803455, CC/AA *vs.* CC, OR=0.89; 95% CI 0.81-0.98; p_{adj}=0.04). In recessive models, the AA genotypes of *RUNX2* (rs13201287, OR=1.18; 95% CI 1.00-1.39; p=0.05), *TGF-βR1* (rs6478974, OR=1.13; 95% CI 1.00-1.28; p=0.05) and *RUNX3* (rs4478762, OR=1.45; 95% CI 0.97-1.27; p=0.07) were also positively associated with an increase in risk, although the associations were borderline significant before adjustment for multiple comparisons. After multiple comparisons adjustment only two SNPs [*RUNX3* (rs906296) and *TGF-β1* (rs4803455)] remained significantly associated with breast cancer.

Table 11a: TGF-β signaling genes and ERα: overall associations with breast cancer risk, *The Breast Cancer Health Disparities Study (Abbreviated table)*

	Controls		Cases					
	N	(%)	N	(%)	OR[a]	(95% CI)		p[b]
RUNX1 (rs7279383)								
CC	3098	73.6	2647	75.1	1.00			0.032 (0.23)
CG/GG	1110	26.4	876	24.9	0.89	(0.80,	0.99)	
RUNX1 (rs8127225)								
TT	2591	61.6	2117	60.1	1.00			0.029 (0.23)
TC/CC	1612	38.4	1407	39.9	1.11	(1.01,	1.22)	
RUNX2 (rs10948238)								
CC/CT	3602	85.6	2950	83.8	1.00			0.028 (0.42)
TT	604	14.4	571	16.2	1.15	(1.01,	1.30)	
RUNX2 (rs13201287)								
GG/GA	3899	92.6	3226	91.5	1.00			0.050 (0.69)
AA	310	7.4	298	8.5	1.18	(1.00,	1.39)	
RUNX3 (rs906296)								
CC	2701	64.2	2143	60.8	1.00			**0.004 (0.03)**
CG/GG	1505	35.8	1379	39.2	1.15	(1.04,	1.26)	
RUNX3 (rs4478762)								
GG/GA	4162	98.9	3470	98.5	1.00			0.066 (0.41)
AA	45	1.1	54	1.5	1.45	(0.97,	1.27)	
TGF-β1 (rs4803455)								
CC	1400	35.3	1193	37.0	1.00			**0.023 (0.04)**
CA/AA	2561	64.7	2032	63.0	0.89	(0.81,	0.98)	
TGF-βR1 (rs6478974)								
TT/TA	3531	83.9	2884	81.9	1.00			0.045 (0.19)
AA	677	16.1	639	18.1	1.13	(1.00,	1.28)	
ERα (rs1801132)								
CC/CG	672	95.3	630	92.2	1.00			0.018 (0.08)
GG	33	4.7	53	7.8	1.72	(1.10,	2.69)	
ERα (rs3798577)								
TT/TC	577	81.8	526	77.0	1.00			0.023 (0.08)
CC	128	18.2	157	23.0	1.36	(1.04,	1.76)	

[a] TGF-β signaling SNPs (n=7733) adjusted for age, study site, and genetic admixture; ERα SNPs adjusted for age and genetic admixture (n=1388)

[b] Wald p-value for 1 df test; Bonferroni-Holm p-value for adjustment for multiple comparisons shown in parentheses

Interaction with Menopausal Status

Several genes (*RUNX1, RUNX3,* and *ERα*) have SNPs that were associated with risk within menopausal strata (Table 12a). Although there were no significant interactions there were associations where risk was divergent among pre- and post-menopausal women. Among post-menopausal women, *RUNX1* (rs2268288, OR_{CC}=1.47; 95% CI 1.03-2.09), *RUNX3* (rs4478762, OR_{AA}=1.71; 95% CI 1.04-2.82) and *ERα* (rs1801132, OR_{GG}=2.14; 95% CI 1.18-3.87) were modestly associated with an increase in risk while pre-menopausal risk was attenuated (OR=0.88-1.33, respectively). The following SNPs were significant for an increase in pre-menopausal risk: *RUNX1* (rs8127225, $OR_{TC/CC}$=1.24; 95% CI 1.06-1.44) and *RUNX3* (rs906296, OR_{AA}=1.33; 95% CI 1.14-1.55); but not in post-menopausal risk (OR=1.04-1.05, respectively).

After adjustment for multiple comparisons within menopause strata, there was one SNP with a p-trend that remained significant in pre-menopausal [*RUNX3* (rs906296)] and post-menopausal [*ERα* (rs1801132)] breast cancer. Associations for all SNPs stratified by menopausal status are found in Table 12b (*See Appendix*).

Table 12a. The association of TGF-β signaling and ERα genes and breast cancer stratified by menopausal status *(Abbreviated table)*

	Pre/Peri Menopause						Post Menopause						
	Controls		Cases				Controls		Cases				
	N	*(%)*	*N*	*(%)*	OR[a]	(95% CI)	*N*	*(%)*	*N*	*(%)*	OR[a]	(95% CI)	*p-int*[b]
RUNX1 (rs2268288)													
TT/TC	1480	97.3	1273	97.5	1.00		2515	97.7	2033	96.6	1.00		0.10 (0.77)
CC	41	2.7	33	2.5	0.88	(0.55, 1.41)	59	2.3	71	3.4	1.47	(1.03, 2.09)	
Wald-p[c]					1.00						0.26		
RUNX1 (rs8127225)													
TT	919	60.6	747	57.2	1.00		1602	62.3	1303	61.9	1.00		0.15 (0.92)
TC/CC	598	39.4	559	42.8	1.24	(1.06, 1.44)	970	37.7	802	38.1	1.04	(0.92, 1.17)	
Wald-p[c]					0.06						1.00		
RUNX3 (rs906296)													
CC	1009	66.3	774	59.3	1.00		1619	63.0	1301	61.9	1.00		0.01 (0.08)
CG/GG	512	33.7	532	40.7	**1.33**	**(1.14, 1.55)**	952	37.0	802	38.1	1.05	(0.93, 1.18)	
Wald-p[c]					**0.002**						1.00		
RUNX3 (rs4478762)													
GG/GA	1504	98.9	1293	99.0	1.00		2546	99.0	2068	98.2	1.00		0.21 (1.00)
AA	16	1.1	13	1.0	0.98	(0.47, 2.05)	27	1.0	37	1.8	1.71	(1.04, 2.82)	
Wald-p[c]					0.95						0.26		
ERα (rs1801132)													
CC/CG	220	94.0	219	92.0	1.00		452	96.2	411	92.4	1.00		0.37 (1.00)
GG	14	6.0	19	8.0	1.36	(0.66, 2.78)	18	3.8	34	7.6	**2.14**	**(1.18, 3.87)**	
Wald-p[c]					0.65						**0.045**		

[a]Odds Ratios adjusted for age, study, and genetic admixture for *RUNX* genes (n=7,506); adjusted for age and genetic admixture for *ERα* (n=1387)

[b]Interaction p-value (gene*menopause); Bonferroni-Holm p-value for multiple comparisons shown in parenthesis

[c] Wald p-value within strata adjusted for multiple comparisons (MC) (Bonferroni-Holm step-down method), **bold** text indicates significance after MC adjustment

Interaction with Proportion Native American Ancestry

Associations were stratified by genetic admixture, based on the distribution of genetic ancestry in the control population: low (0-28%), moderate (29-70%), and high (71-100%) proportion NA ancestry (Table 13a). There was a significant interaction with genetic admixture and *RUNX1* (rs7279383, p_{adj}=0.04) after adjustment for multiple comparisons. In the dominant model of rs7279383 (CG/GG *vs.* CC) results were divergent between strata; individuals in the 71-100% admixture strata had a significant increase in risk ($OR_{CG/GG}$ =1.75 95% CI 1.17-2.63) while those in the 0-28% and 29-70% had a reduced risk [($OR_{CG/GG}$ =0.87 95% CI 0.76-1.00) and ($OR_{CG/GG}$ =0.82 95% CI 0.69-0.97), respectively]. There were no other significant interactions, although a few SNPs were associated in genetic admixture strata. Individuals in the 29-70% strata had a significantly higher risk with *RUNX1* (rs8127225, $OR_{TC/CC}$=1.19); *RUNX2* (rs6930053, OR_{TT}=1.29); and *TGF-β1* (rs1800469, OR_{TT}=1.29); while there was a null or inverse association within 0-28% and 29-70% strata for these genotypes. The AA genotype of *RUNX3* (rs4478762) risk was >2 fold for the 0-28% strata, while the 29-70% and 71-100% were not positively associated. Risk was similar between low and high admixture strata for RUNX2 [(rs10948238, OR_{TT}=1.27 and 1.34) and (rs13201287, OR_{AA}=1.39 and 1.36)]; although it was null for the 29-71% admixture group. Risk was similar for RUNX3 (rs906296) for 29-71% and 71-100% admixture groups ($OR_{CG/GG}$=1.23 and 1.24), while it was towards the null for 0-28% ($OR_{CG/GG}$=1.05). When adjusting for multiple comparisons by admixture strata, two SNPs remained positively associated with moderate admixture and one with high admixture. Table 13b shows associations for all SNPs stratified by genetic admixture (*See Appendix*).

Table 13a. The association of TGF-β signaling genes and breast cancer stratified by proportion Native American ancestry *(Abbreviated table)*

Gene (SNP)	Genotype	Low (0-28%)		Moderate (29-70%)		High (71-100%)		p-int [b]
		Case/Control	OR (95% CI) [a]	Case/Control	OR (95% CI) [a]	Case/Control	OR (95% CI) [a]	
RUNX1 (rs7279383)								
	CC	1165/1229	1.00	1092/1285	1.00	390/584	1.00	**0.004 (0.04)**
	CG/GG	530/641	0.87 (0.76, 1.00)	288/419	**0.82 (0.69, 0.97)**	58/50	**1.75 (1.17, 2.63)**	
Wald-p [c]			0.41		**0.14**		**0.05**	
RUNX1 (rs8127225)								
	TT	1229/1380	1.00	693/931	1.00	195/280	1.00	0.48 (1.00)
	TC/CC	467/490	1.05 (0.90, 1.22)	687/770	**1.19 (1.03, 1.38)**	253/352	1.01 (0.79, 1.29)	
Wald-p [c]			1.00		**0.12**		1.00	
RUNX2 (rs6930053)								
	CC	628/636	1.00	625/813	1.00	265/371	1.00	0.15 (1.00)
	CT	808/944	0.87 (0.75, 1.00)	596/731	1.06 (0.91, 1.23)	158/224	0.98 (0.75, 1.27)	
	TT	259/291	0.91 (0.74, 1.11)	159/160	1.29 (1.01, 1.65)	25/39	0.90 (0.53, 1.54)	
p-trend [c]			1.00		0.57		1.00	
RUNX2 (rs10948238)								
	CC/CT	1398/1603	1.00	1171/1440	1.00	381/559	1.00	0.13 (1.00)
	TT	295/267	**1.27 (1.06, 1.52)**	209/262	0.98 (0.81, 1.20)	67/75	1.34 (0.93, 1.91)	
Wald-p [c]			**0.12**		1.00		1.00	
RUNX2 (rs13201287)								
	GG/GA	1568/1767	1.00	1269/1562	1.00	389/570	1.00	0.10 (0.88)
	AA	128/104	**1.39 (1.06, 1.82)**	111/142	0.95 (0.73, 1.23)	59/64	1.36 (0.93, 1.99)	
Wald-p [c]			**0.17**		1.00		1.00	
RUNX3 (rs906296)								
	CC	987/1113	1.00	873/1155	1.00	283/433	1.00	0.22 (1.00)
	CG/GG	707/758	1.05 (0.92, 1.20)	507/546	**1.23 (1.06, 1.43)**	165/201	1.24 (0.96, 1.61)	

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			p-int [b]
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		
Wald-p [c]			0.94			**0.04**			0.61		
RUNX3 (rs4478762)											
	GG/GA	1670/1856	1.00		1358/1681	1.00		442/625	1.00		0.48 (1.00)
	AA	26/15	2.01	(1.06, 3.81)	22/22	1.24	(0.68, 2.25)	6/8	1.09	(0.38, 3.18)	
Wald-p [c]			0.20			1.00			1.00		
TGF-β1 (rs1800469)											
	CC	772/866	1.00		379/516	1.00		102/141	1.00		0.42 (0.84)
	CT	692/759	1.01	(0.88, 1.17)	692/854	**1.11**	**(0.94, 1.31)**	206/299	0.95	(0.69, 1.30)	
	TT	190/216	0.96	(0.77, 1.20)	295/313	**1.29**	**(1.04, 1.58)**	133/193	0.92	(0.65, 1.30)	
p-trend [c]			1.00			**0.04**			0.94		
TGF-β1 (rs4803455)											
	CC	444/478	1.00		509/604	1.00		240/318	1.00		0.63 (0.84)
	CA/AA	1173/1339	0.95	(0.82, 1.11)	670/932	**0.86**	**(0.73, 1.00)**	189/290	0.91	(0.71, 1.17)	
Wald-p [c]			1.00			**0.06**			0.94		

[a] Odds Ratios adjusted for age and study

[b] Interaction p-value (SNP*admixture); Bonferroni-Holm p-value for multiple comparisons shown in parenthesis; **bold** text indicates significance after MC adjustment

[c] Wald (or trend) p-value within strata adjusted for multiple comparisons (MC) by admixture strata (Bonferroni-Holm step-down method), **bold** text indicates significance ($p \leq 0.05$) or suggestive of an association ($p \leq 0.15$) after MC adjustment

Table 14 shows *ERα* SNPs and breast cancer by genetic admixture strata (≤28%, 29-70%). No *ERα* SNPs were found to significantly interact with genetic admixture. However, within the ≤28% NA ancestry strata, there was a trend for *ERα* (rs3798577, OR_{CC}=1.43 95% CI 1.06-1.52, p=0.02), while the risk in 29-70% NA ancestry strata was attenuated and not significant (OR_{CC}=1.11 95% CI 0.65-1.91). However, this trend did not remain significant after adjustment for multiple comparisons. Although not significant, findings were comparable between strata for *ERα* (rs1801132, rs2046210 and rs6913578) and divergent for *ERα* [rs851984, ≤28% (OR_{AA}=1.34 95% CI 0.65-1.91); 29-70% (OR_{AA}=0.83 95% CI 0.46-1.50)].

Table 14. The Association ERα SNPs and breast cancer stratified by proportion Native American ancestry

	Low (0-28%)						Moderate to High (29-100%)						
	Cases (n=491)		Controls (n=506)				Cases (n=192)		Controls (n=199)				
	N	(%)	N	(%)	OR[a]	(95% CI)	N	(%)	N	(%)	OR[a]	(95% CI)	*p-int*[b]
rs1801132													
CC/CG	458	93.3	486	96.1	1.00		172	89.6	186	93.5	1.00		0.92 (1.00)
GG	33	6.7	20	3.9	1.74	(0.98-3.08)	20	10.4	13	6.5	1.80	(0.86-3.75)	
Wald-p [c]					0.06 (0.22)						0.12 (0.55)		
rs3798577													
TT/TC	368	74.9	410	81.0	1.00		158	82.3	167	83.9	1.00		0.41 (1.00)
CC	123	25.1	96	19.0	1.43	(1.05-1.93)	34	17.7	32	16.1	1.07	(0.62-1.85)	
Wald-p [c]					0.02 (0.09)						0.80 (1.00)		
rs2046210													
GG	183	37.3	212	41.9	1.00		98	51.0	110	55.3	1.00		0.93 (1.00)
GA/AA	308	62.7	294	58.1	1.24	(0.96-1.60)	94	49.0	89	44.7	1.17	(0.78-1.75)	
Wald-p [c]					0.11 (0.29)						0.45 (1.00)		
rs851984													
GG	180	36.7	202	39.9	1.00		71	36.9	70	35.1	1.00		0.40 (1.00)
GA	231	47.0	237	46.8	1.09	(0.83-1.43)	91	47.4	93	46.4	0.95	(0.61-1.48)	
AA	80	16.3	67	13.2	1.33	(0.90-1.94)	30	15.6	36	18.5	0.85	(0.47-1.53)	
p-trend[c]					0.16 (0.29)						0.59 (1.00)		
rs6913578													
AA	205	41.8	229	45.3	1.00		103	53.6	117	58.8	1.00		0.90 (1.00)
AC	223	45.4	218	43.1	1.16	(0.89-1.52)	75	39.1	71	35.7	1.20	(0.78-1.83)	
CC	63	12.8	59	11.7	1.22	(0.81-1.82)	14	7.3	11	5.5	1.45	(0.63-3.37)	
p-trend[c]					0.23 (0.29)						0.27 (0.97)		

[a] Odds Ratios and 95% CIs are adjusted for age

[b] Interaction p-value (SNP*genetic admixture); Bonferroni-Holm p-value adjusted for multiple comparisons shown in parentheses

[c] Trend (additive) or Wald (dominant/recessive) p-value within strata adjusted for multiple comparisons (Bonferroni-Holm step-down method) shown in parentheses

Association with Breast Cancer, Defined by ER/PR expression Status

Using controls (n=3,215) as the referent group, *TGF-β* signaling SNPs were assessed for the association with risk of breast cancer defined ER/PR phenotypes (cases, n=1,963) (Table 15a). When evaluating the Wald p-value for overall models, there were five SNPs [*RUNX1* (rs7279123, p=0.01), *RUNX2* (rs9463090, p=0.01), *RUNX2* (rs12333172, p=0.01), *RUNX3* (rs2236850, p=0.03; rs7517302, p=0.005) and *TGF-βR1* (rs10733710, p=0.05)] significant at the 0.05 level; however, only one was significant after multiple comparisons (*RUNX3*, rs7517302). Several SNPs were associated with breast cancer within specific ER/PR tumor phenotype strata. Variants in *RUNX1* (1 SNP), *RUNX3* (2 SNPs), and *TGF-βR1* (1 SNP) were associated with ER+/PR+ tumors (OR between 0.78 and 1.90); in *RUNX1* (2 SNPs) and *TGF-βR1* (2 SNPs) were associated with ER+/PR- tumors (OR between 0.44 and 3.55); in *RUNX3* (2 SNPs) were associated with ER-/PR+ tumors (OR between 2.52 and 2.88); and in *RUNX1* (1 SNP), *RUNX2* (5 SNPs), and *RUNX3* (2 SNPs) were associated with ER-/PR- tumors (OR between 0.71-2.31).

The Wald-p for the overall dominant or recessive models remained significant for one SNP (*RUNX3* (rs7517302), p_{adj}=0.04. The *p-trend* for 5 SNPs within ER/PR tumor phenotype strata [ER+/PR+ (n=0); ER+/PR- (*RUNX1*, rs7279123); ER-/PR+ (*RUNX3*, rs2236850); ER-/PR- (RUNX2, rs9463090, rs12333172; RUNX3, rs7517302)] remained significant at 0.05 level after adjustment for multiple comparisons (Table 15a). Results for all associations with breast cancer defined by ER/PR status are found in Table 15b. (*See Appendix*)

Table 15a. The Association of TGF-β signaling genes and breast cancer defined by ER/PR status *(Abbreviated table)*

	Controls[a]	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
	N	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)	p[c]
***RUNX1* (rs7279383)**														
CC	2249	927	1.00		167	1.00		32	1.00		306	1.00		0.13 (0.74)
CG/GG	965	351	0.87	(0.75, 1.00)	66	0.92	(0.69, 1.24)	13	0.99	(0.51, 1.91)	100	0.77	(0.60, 0.98)	
***RUNX1* (rs8127225)**														
TT	2132	821	1.00		137	1.00		28	1.00		262	1.00		0.10 (0.65)
TC/CC	1078	458	1.14	(0.99, 1.31)	96	1.40	(1.06, 1.85)	17	1.08	(0.58, 2.02)	144	1.04	(0.83, 1.30)	
***RUNX1* (rs7279123)**														
CC	1935	730	1.00		160	1.00		31	1.00		249	1.00		0.01 (0.06)
CT	1083	466	1.12	(0.97, 1.29)	65	**0.72**	**(0.53, 0.97)**	14	0.83	(0.44, 1.57)	135	0.98	(0.78, 1.22)	
TT	184	79	1.10	(0.83, 1.46)	7	**0.44**	**(0.21, 0.96)**	0	0.00	(0.00, 0.00)	20	0.88	(0.55, 1.43)	
***RUNX2* (rs1321075)**														
CC	1918	741	1.00		139	1.00		23	1.00		240	1.00		0.44 (1.00)
CA	1102	456	1.11	(0.97, 1.28)	80	1.00	(0.75, 1.34)	19	1.33	(0.71, 2.51)	127	0.90	(0.72, 1.14)	
AA	195	81	1.17	(0.89, 1.55)	14	1.00	(0.56, 1.80)	3	1.18	(0.34, 4.09)	39	1.56	(1.06, 2.29)	
***RUNX2* (rs9463090)**														
GG	2051	818	1.00		155	1.00		25	1.00		238	1.00		0.01 (0.09)
GA	1016	410	1.02	(0.88, 1.17)	70	0.93	(0.70, 1.25)	17	1.34	(0.72, 2.50)	130	**1.10**	**(0.87, 1.38)**	
AA	144	49	0.83	(0.59, 1.16)	8	0.72	(0.35, 1.50)	3	1.81	(0.54, 6.10)	38	**2.31**	**(1.57, 3.39)**	
***RUNX2* (rs6930053)**														
CC	1280	493	1.00		86	1.00		18	1.00		183	1.00		0.16 (1.00)
CT	1516	598	1.01	(0.88, 1.16)	118	1.16	(0.87, 1.54)	20	0.99	(0.52, 1.89)	181	0.85	(0.68, 1.06)	
TT	419	187	1.13	(0.92, 1.38)	29	1.03	(0.66, 1.59)	7	1.28	(0.53, 3.11)	42	0.71	(0.50, 1.02)	
***RUNX2* (rs12209785)**														
AA/AG	3024	1196	1.00		220	1.00		42	1.00		372	1.00		0.35 (1.00)
GG	190	82	1.10	(0.84, 1.44)	13	0.92	(0.52, 1.65)	3	1.20	(0.37, 3.93)	34	1.49	(1.01, 2.18)	
***RUNX2* (rs10948238)**														

	Controls[a]	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
	N	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)	p[c]
CC/CT	2742	1072	1.00		199	1.00		36	1.00		328	1.00		0.12 (1.00)
TT	470	205	1.11	(0.93, 1.33)	33	0.97	(0.66, 1.41)	9	1.47	(0.70, 3.09)	78	1.39	(1.07, 1.82)	
***RUNX2* (rs12333172)**														
CC/CT	3109	1238	1.00		222	1.00		43	1.00		378	1.00		0.01 (0.10)
TT	106	40	0.92	(0.63, 1.33)	11	1.44	(0.76, 2.73)	2	1.36	(0.32, 5.72)	28	**2.12**	**(1.37, 3.27)**	
***RUNX3* (rs2236850)**														
TT	1073	401	1.00		77	1.00		11	1.00		105	1.00		0.03 (0.24)
TC/CC	2135	877	1.09	(0.95, 1.26)	156	1.01	(0.76, 1.35)	33	**1.51**	**(0.76, 3.00)**	299	1.43	(1.13, 1.80)	
***RUNX3* (rs9438876)**														
AA	871	325	1.00		63	1.00		8	1.00		100	1.00		0.1 (0.59)
AG	1527	634	1.10	(0.94, 1.28)	120	1.08	(0.79, 1.49)	19	1.37	(0.60, 3.15)	192	1.09	(0.84, 1.41)	
GG	817	319	1.01	(0.85, 1.22)	50	0.83	(0.57, 1.23)	18	2.52	(1.08, 5.87)	113	1.21	(0.91, 1.62)	
***RUNX3* (rs7517302)**														
TT	1123	431	1.00		91	1.00		10	1.00		114	1.00		**.005 (0.04)**
TC	1541	604	1.01	(0.88, 1.17)	100	0.80	(0.59, 1.07)	22	1.68	(0.79, 3.58)	210	**1.36**	**(1.07, 1.74)**	
CC	548	242	1.14	(0.95, 1.38)	41	0.93	(0.63, 1.37)	13	2.80	(1.22, 6.46)	82	**1.51**	**(1.11, 2.04)**	
***RUNX3* (rs906296)**														
CC	2009	748	1.00		134	1.00		27	1.00		247	1.00		0.13 (0.59)
CG/GG	1203	530	1.18	(1.03, 1.34)	98	1.22	(0.93, 1.61)	18	1.16	(0.63, 2.12)	159	1.09	(0.88, 1.35)	
***RUNX3* (rs4478762)**														
GG/GA	3178	1253	1.00		232	1.00		45	1.00		400	1.00		0.11 (0.59)
AA	35	26	1.90	(1.14, 3.18)	1	0.39	(0.05, 2.88)	0	0.00	(0.00, 0.00)	6	1.35	(0.56, 3.23)	
***TGF-βR1* (rs11568785)**														
AA/AG	3199	1272	1.00		229	1.00		44	1.00		405	1.00		0.09 (0.29)
GG	16	7	1.08	(0.44, 2.63)	4	3.55	(1.17, 10.8)	1	5.70	(0.73, 44.6)	1	0.55	(0.07, 4.21)	
***TGF-βR1* (rs10733710)**														
GG/GA	2938	1195	1.00		222	1.00		41	1.00		380	1.00		0.05 (0.23)
AA	276	84	0.78	(0.60, 1.00)	11	0.53	(0.28, 0.99)	4	0.92	(0.32, 2.61)	25	0.67	(0.44, 1.03)	

[a] ER/PR data were compared with 3,214 controls from sites where cases have ER/PR data Mexico data is excluded because they do not have data for ER/PR status (n=1,810)

[b] Odds ratios (OR) and 95% confidence intervals (CI) adjusted for age, study, and genetic admixture; Bonferroni-Holms p-value adjustment for multiple comparisons shown in parentheses. Note: OR in **bold text** indicates significance remained for wald-p (dominant/recessive) or p-trend (additive) after adjustment for multiple comparisons

Table 16 reports the association of *ERα* SNPs and breast cancer defined by ER+/- tumor phenotype (sample size limits evaluation by ER/PR status combined) adjusting for age and genetic admixture. Participants with available data on ER status and genetic admixture were used in analysis [n=1,135 (controls=705, referent; cases=430)]. Risk was similar for ER+ and ER- tumors for rs3798577 (OR_{CC}=1.35 95% CI 0.98-1.36 and OR_{CC}=1.41 95% CI 0.84-2.36, respectively. Risk was moderately higher among ER- tumors for rs1801132 (OR_{GG}=2.04) and rs851984 (OR_{AA}=1.54) compared to ER+ tumors (OR=1.56 and 1.04, respectively). Risk was slightly higher among ER+ tumors for rs2046210 (OR_{CC}=1.23) and rs6913578 (OR_{CC}=1.30) compared to ER- tumors (OR=1.10 and 1.02, respectively). None of these associations were significant for the overall model or within SNP (genotype)-tumor phenotype strata before or after multiple comparisons.

Table 16: The Association of ERα SNPs and breast cancer defined by ER status [a]

		Controls		ER+				ER-				
SNP	***Genotype***	**N**	**(%)**	**N**	**(%)**	**OR[b]**	**(95% CI)**	**N**	**(%)**	**OR[b]**	**(95% CI)**	**p[c]**
rs1801132												
	CC/CG	672	95.3	313	93.2	1.00		85	90.4	1.00		0.11 (0.51)
	GG	33	4.7	23	6.8	1.56	(0.90-2.72)	9	9.6	2.04	(0.94-4.43)	
	Wald-p [d]					0.11 (0.40)				0.07 (0.32)		
rs3798577												
	TT/TC	577	81.8	259	77.1	1.00		71	75.5	1.00		0.12 (0.51)
	CC	128	18.2	77	22.9	1.35	(0.98-1.86)	23	24.5	1.41	(0.84-2.36)	
	Wald-p [d]					0.07 (0.32)				0.19 (0.65)		
rs2046210												
	GG	322	45.7	135	40.2	1.00		41	43.6	1.00		0.32 (0.83)
	GA/AA	383	54.3	201	59.8	1.23	(0.94-1.60)	53	56.4	1.10	(0.71-1.72)	
	Wald-p [d]					0.13 (0.40)				0.66 (1.00)		
rs851984												
	GG	272	38.6	124	36.9	1.00		31	32.9	1.00		0.71 (0.97)
	GA	330	46.8	164	48.8	1.10	(0.83-1.46)	45	47.9	1.19	(0.73-1.94)	
	AA	103	14.6	48	14.3	1.04	(0.70-1.56)	18	19.2	1.54	(0.82-2.88)	
	p-trend [d]					0.70 (0.70)				0.18 (0.65)		
rs6913578												
	AA	346	49.1	145	43.2	1.00		45	47.9	1.00		0.61 (0.97)
	AC	289	41.0	152	45.2	1.23	(0.93-1.63)	40	42.6	1.08	(0.69-1.71)	
	CC	70	9.9	39	11.6	1.30	(0.83-2.02)	9	9.6	1.02	(0.47-2.20)	
	p-trend [d]					0.12 (0.40)				0.83 (1.00)		

[a] ER data is available for 1135 subjects (430 cases (ER+=336; ER-=94) and is compared to 705 controls)
[b] OR (odds ratios) and 95% confidence interval (CI) adjusted for age and genetic admixture
[c] Wald-p for over all model; Bonferroni-Holm multiple comparison adjustments shown in parentheses
[d] Wald p for dominant/recessive models, p-trend for additive models within ER+ and ER- strata, Bonferroni-Holm multiple comparison adjustments shown in parentheses

Association with Breast Cancer, Defined by ER Expression Status: Stratified by Menopause

TGF-β signaling SNPs were also evaluated for their association with breast cancer, defined by ER status, stratified by menopausal status (sample size limits evaluation by ER/PR status combined, as well as associations with *ERα* SNPs). Table 17a shows SNPs associated with ER+/ER- breast cancer that differ by menopausal status after adjustment for multiple comparisons. Among pre-menopausal women, the association was suggestive for an increase in risk of ER+ breast cancer with *RUNX3* (rs906296, $OR_{CG/GG}$=1.33, p=0.06), while an inverse association was observed with *RUNX2* (rs598953, OR_{AA}=0.61, p=0.07). These associations were divergent from the risk of ER- tumors, however not significant. A suggestive inverse association was observed for both ER+ and ER- tumors with *TGF-βR1* (rs10733710, OR_{AA}=0.58 and 0.45, respectively). The findings for these particular SNPs were not similar in post-menopausal women.

Among post-menopausal women, there were two SNPs significantly suggestively associated with an increase in risk of ER+ breast cancer: *RUNX3* (rs4478762), with a >2-fold increase in risk (OR_{AA}=2.08, p=0.13), which was not observed for ER- tumors. In contrast, *TGF-β1* (rs4803455) was associated with a reduced risk of ER+ tumors (OR_{AA}=0.82, p=0.13); the trend was similar for ER- tumor, although not significant.

With regards to risk of ER- tumors, one association suggestive of a reduced risk was observed with *RUNX1* (rs7279383, OR_{GG}=0.76, p=0.16), while risk for ER+ tumors with was null (OR_{GG}=1.01). *RUNX2* (rs12333172) was associated with a suggestive increase in risk of ER- tumors (OR_{AA}=1.97, p=0.13), while *RUNX1* (rs2268288) and

RUNX2 (rs12333172) were significantly associated with a modest increase in risk for ER- tumors (OR_{CC}=2.47, p=0.04 and OR_{TT}=2.25, p=0.04, respectfully), which was not observed for ER+ tumors. There were no significant interactions between SNPs and menopausal status for ER+ or ER- breast cancer; i.e. the outcome did not differ within menopausal status and SNP categories. Associations for all SNPs and breast cancer defined by ER, stratified by menopausal status, can be found in Table 17b (*See Appendix*).

Association with Breast Cancer, Defined by ER Expression Status:
Stratified by Proportion Native American Ancestry

Table 18a shows SNPs associated with ER+/ER- breast cancer that differ by proportion Native American ancestry (low, moderate/high) after adjustment for multiple comparisons. (Note: moderate to high ancestry were grouped together to increase power). Among <u>low proportion</u> (0-28%) of Native American ancestry, *RUNX2* (rs9463090 and rs10948238) was associated with an increase in risk of ER- tumors (OR_{AA}=2.03 and OR_{TT}=1.68, respectfully). There were a larger number of significant associations for those with <u>moderate to high</u> (29-100%) Native American ancestry: *RUNX1* (rs8127225) and *RUNX3* (rs906296) were both positively associated with ER+ tumors; while RUNX3 (rs2236850) was associated with ER- tumors. Regardless of ER status, *TGF-βRI* (rs10733710) had an inverse association with breast cancer risk, although it was stronger for ER- *vs.* ER+ tumors (OR_{AA}=0.48 *vs.* OR_{AA}=0.64). Associations for all SNPs and breast cancer defined by ER, stratified by proportion Native American ancestry, can be found in Table 18b (*See Appendix*).

Table 17a. TGF-β signaling SNPs and breast cancer defined by ER status, stratified by menopausal status (*Abbreviated table*)

	Pre-menopausal								Post-menopausal								
	Controls[a]	ER+				ER-				ER+				ER-			
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
RUNX1 (rs7279383)																		
CC	2166	351	1.00			145	1.00			694	1.00			175	1.00			1.00
CG	864	130	0.88	(0.69,	1.13)	51	0.86	(0.61,	1.23)	242	0.87	(0.72,	1.03)	45	**0.65**	**(0.46,**	**0.92)**	
GG	71	11	0.92	(0.44,	1.90)	3	0.64	(0.19,	2.15)	23	1.01	(0.61,	1.68)	4	**0.76**	**(0.27,**	**2.13)**	
p-trend			1.00				1.00				1.00				**0.16**			
RUNX1 (rs2268288)																		
TT/TC	3010	477	1.00			194	1.00			925	1.00			211	1.00			0.59
CC	91	15	0.83	(0.45,	1.53)	5	0.70	(0.27,	1.80)	34	1.37	(0.88,	2.13)	13	**2.47**	**(1.31,**	**4.65)**	
Wald-p			1.00				1.00				1.00				**0.04**			
RUNX2 (rs9463090)																		
GG	1976	319	1.00			115	1.00			614	1.00			131	1.00			1.00
GA	978	152	0.89	(0.71,	1.12)	66	1.05	(0.76,	1.47)	309	1.08	(0.91,	1.28)	73	**1.19**	**(0.88,**	**1.62)**	
AA	143	21	1.04	(0.60,	1.79)	18	2.51	(1.39,	4.54)	35	0.71	(0.48,	1.06)	20	**1.97**	**(1.17,**	**3.30)**	
p-trend			1.00				0.34				1.00				**0.13**			
RUNX2 (rs12333172)																		
CC/CT	2997	472	1.00			188	1.00			927	1.00			207	1.00			1.00
TT	104	20	1.20	(0.68,	2.10)	11	1.67	(0.83,	3.34)	31	0.90	(0.59,	1.40)	17	**2.25**	**(1.29,**	**3.93)**	
Wald-p			1.00				0.90				1.00				**0.04**			
RUNX2 (rs598953)																		
TT	1129	203	1.00			71	1.00			354	1.00			83	1.00			0.62
TA	1489	233	**0.85**	**(0.68,**	**1.07)**	97	1.02	(0.73,	1.42)	456	0.99	(0.84,	1.18)	118	1.09	(0.81,	1.47)	
AA	483	56	**0.61**	**(0.43,**	**0.87)**	31	0.97	(0.61,	1.54)	149	1.02	(0.81,	1.29)	23	0.64	(0.39,	1.04)	
p-trend			**0.07**				1.00				1.00				1.00			

	Pre-menopausal								Post-menopausal								
	Controls[a]	**ER+**				**ER-**				**ER+**				**ER-**				
	N	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**p-int[c]**
RUNX3 (rs906296)																		
CC	1936	276	1.00			125	1.00			570	1.00			132	1.00			1.00
CG/GG	1162	216	**1.33**	**(1.07,**	**1.65)**	74	1.02	(0.74,	1.39)	387	1.11	(0.95,	1.30)	92	1.16	(0.87,	1.54)	
Wald-p			**0.06**				1.00				0.76				0.78			
RUNX3 (rs4478762)																		
GG/GA	3066	487	1.00			197	1.00			939	1.00			221	1.00			1.00
AA	33	5	0.94	(0.33,	2.71)	2	0.84	(0.19,	3.78)	20	**2.08**	**(1.12,**	**3.86)**	3	1.21	(0.36,	4.12)	
Wald-p			1.00				1.00				**0.13**				0.78			
TGF-β1 (rs4803455)																		
CC	898	134	1.00			59	1.00			291	1.00			64	1.00			0.48
CA	1414	200	0.98	(0.75,	1.27)	74	0.85	(0.58,	1.23)	406	**0.83**	**(0.69,**	**1.00)**	92	0.86	(0.62,	1.21)	
AA	560	95	1.14	(0.82,	1.59)	32	0.94	(0.58,	1.53)	162	**0.82**	**(0.65,**	**1.04)**	32	0.76	(0.49,	1.19)	
p-trend			0.99				0.92				**0.13**				0.27			
TGF-βRI (rs10733710)																		
GG/GA	2835	466	1.00			190	1.00			895	1.00			208	1.00			0.98
AA	265	26	**0.58**	**(0.37,**	**0.92)**	9	**0.45**	**(0.22,**	**0.91)**	64	0.82	(0.61,	1.11)	15	0.78	(0.45,	1.35)	
Wald-p			**0.08**				**0.10**				0.61				1.00			

[a] ER/PR data were compared with 3,214 controls from sites where cases have ER/PR data Mexico data is excluded because they do not have data for ER/PR status (n=1,810)

[b] Odds ratios (OR) and 95% confidence intervals (CI) adjusted for age, study, and genetic admixture; OR in **bold text** indicates significance remained for p-trend (per-allele) or Wald-p after adjustment for multiple comparisons

[c] p-value for interaction term (SNP*menopause) for ER+ or ER- breast cancer as the outcome; Bonferroni-Holms p-value adjustment for multiple comparisons shown

Table 18a. TGF-β signaling and breast cancer defined by ER status, stratified by Native American ancestry (*Abbreviated table*)

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]	ER+			ER-				ER+				ER-				
	N	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
RUNX1 (rs8127225)																	
TT	2132	682	1.00		174	1.00			276	1.00			116	1.00			0.38
TC/CC	1078	251	1.02	(0.86, 1.22)	70	1.11	(0.82, 1.50)		303	**1.39**	**(1.14, 1.69)**		91	0.98	(0.73, 1.32)		
Wald-p			1.00			1.00				**0.008**				1.00			
RUNX2 (rs9463090)																	
GG	2051	586	1.00		131	1.00			387	1.00			132	1.00			0.98
GA	1016	304	1.01	(0.85, 1.19)	89	**1.29**	**(0.97, 1.72)**		176	1.00	(0.81, 1.24)		58	0.94	(0.68, 1.32)		
AA	144	42	0.82	(0.56, 1.20)	24	**2.03**	**(1.24, 3.30)**		15	0.80	(0.44, 1.45)		17	2.73	(1.51, 4.95)		
p-trend			1.00			**0.03**				0.77				0.73			
RUNX2 (rs10948238)																	
CC/CT	2742	775	1.00		191	1.00			496	1.00			173	1.00			0.78
TT	470	155	1.20	(0.96, 1.49)	53	**1.68**	**(1.20, 2.34)**		83	0.94	(0.72, 1.25)		34	1.11	(0.74, 1.65)		
Wald-p			1.00			**0.03**				0.77				1.00			
RUNX3 (rs2236850)																	
TT	1073	296	1.00		63	1.00			182	1.00			53	1.00			1.00
TC	1526	451	1.03	(0.86, 1.24)	126	1.31	(0.95, 1.81)		281	1.13	(0.90, 1.41)		105	**1.47**	**(1.03, 2.09)**		
CC	609	185	1.00	(0.80, 1.25)	54	1.32	(0.89, 1.94)		116	1.26	(0.95, 1.67)		47	**1.79**	**(1.17, 2.74)**		
p-trend			0.96			0.55				0.49				**0.03**			
RUNX3 (rs906296)																	
CC	2009	535	1.00		140	1.00			347	1.00			134	1.00			1.00
CG/GG	1203	396	1.10	(0.94, 1.29)	104	1.08	(0.83, 1.42)		232	**1.33**	**(1.08, 1.62)**		73	1.10	(0.81, 1.50)		
Wald-p			0.72			1.00				**0.04**				1.00			

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]	ER+				ER-				ER+				ER-			
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
TGF-βR1 (rs10733710)																		
GG/GA	2938	890	1.00			229	1.00			527	1.00			192	1.00			0.18
AA	276	43	0.90	(0.62,	1.31)	14	1.20	(0.67,	2.15)	52	**0.64**	**(0.46,**	**0.89)**	15	**0.48**	**(0.28,**	**0.84)**	
Wald-p			1.00				1.00				**0.04**				**0.05**			

[a] ER data were compared with 3,214 controls from sites where cases have ER data Mexico data is excluded because they do not have data for ER status (n=1,810)

[b] Odds ratios (OR) and 95% confidence intervals (CI) adjusted for age, study, and genetic admixture; OR in **bold text** indicates significance ($p \leq 0.05$) or suggestive of an association ($p \leq 0.15$) for p-trend (per-allele) or Wald-p after adjustment for multiple comparisons

[c] p-value for interaction term (SNP*admixture) for ER+ or ER- breast cancer as the outcome; Bonferroni-Holms p-value adjustment for multiple comparisons shown

SNP-SNP Interaction and Association with Breast Cancer

SNP-SNP interactions were evaluated between SNPs with a potentially meaningful overall association ($p<0.15$) with breast cancer. A combination of two SNPs (dominant/recessive mode) resulted in a 4 category variable, the referent as the combination of the referent groups for both SNPs being evaluated. Table 19 shows results for interactions between *RUNX1* (rs7279383, rs2268288, rs8127225, rs7279123)**TGF-β1* (rs4803455); and *RUNX1*TGF-βRI* (rs6478974). There were no significant p-values for the interaction term in the model for each pair of SNPs. There were few significant associations evaluating the combined risk. (Note: minor alleles are denoted with **bolded** text). There was an inverse association for rs7279383*rs4803455 (OR=0.79 95% CI 0.68-0.92); where either one/two copies (C**G**/**GG**-C**A**/**AA**) of the minor allele lowered risk. Others associated with an inverse association include: rs7279383*rs6478974 (C**G**/**GG**-TT/T**A**, OR=0.87 95% CI 0.79-0.98); rs2268288*rs4803455 (TT/T**C**-C**A**/**AA**, OR=0.89 95% CI 0.81-0.99); and rs7279123*4803455 (C**T**/**TT**-C**A**/**AA**, OR=0.81 95% CI 0.71-0.93). A modest increase in risk was observed when four minor alleles are present for rs2268288*rs4803455 (**CC-AA,** OR=1.83 95% CI 1.02-3.27), while there was a slight increase for rs8127225*rs4803455 [(TT-**AA,** OR=1.14 95% CI 1.03-1.26 and T**C**/**CC**-TT/T**A,** OR=1.20 95% CI 1.03-1.39)], which appear to be driven by the *TGF-β1* minor alleles. All other combinations of *RUNX1*TGF-β1* and *RUNX1*TGF-βRI* SNPs were not significantly associated with risk among categories of combined genotypes.

Table 19. Interactions between RUNX1, TGF-β, TGF-βRI genes and breast cancer

			Combined Risk		
Gene (SNP)[a]	Gene (SNP)[a]	Cases/Controls	OR (95% CI)[b]	p^c	*p-int*[d]
RUNX1 (rs7279383)	*TGF-β1* (rs4803455)			0.02	0.86
CC	CC	914/1060	1.00 (REF)		
CC	C**A**/**AA**	1508/1854	0.90 (0.80-1.00)		
C**G**/**GG**	CC	279/340	0.90 (0.75-1.08)		
C**G**/**GG**	C**A**/**AA**	523/707	**0.79 (0.68-0.92)**		
RUNX1 (rs7279383)	*TGF-βR1* (rs6478974)			0.03	0.45
CC	TT/T**A**	2191/2618	1.00 (REF)		
CC	**AA**	456/480	1.10 (0.96-1.27)		
C**G**/**GG**	TT/T**A**	692/912	**0.87 (0.79-0.98)**		
C**G**/**GG**	**AA**	183/197	1.06 (0.86-1.32)		
RUNX1 (rs2268288)	*TGF-β1* (rs4803455)			0.08	0.94
TT/T**C**	CC	1167/1377	1.00 (REF)		
TT/T**C**	C**A**/**AA**	1959/2487	**0.89 (0.81-0.99)**		
CC	CC	25/23	1.19 (0.67-2.11)		
CC	C**A**/**AA**	73/74	1.08 (0.77-1.51)		
RUNX1 (rs2268288)	*TGF-βR1* (rs6478974)			0.06	0.28
TT/T**C**	TT/T**A**	2804/3448	1.00 (REF)		
TT/T**C**	**AA**	610/658	1.12 (0.99-1.26)		
CC	TT/T**A**	79/83	1.13 (0.83-1.55)		
CC	**AA**	29/19	**1.83 (1.02-3.27)**		
RUNX1 (rs8127225)	*TGF-β1* (rs4803455)			0.02	0.64
TT	CC	678/806	1.00 (REF)		
TT	C**A**/**AA**	1269/1645	0.89 (0.77-1.00)		
T**C**/**CC**	CC	515/591	1.09 (0.93-1.27)		
T**C**/**CC**	C**A**/**AA**	763/913	1.00 (0.87-1.15)		
RUNX1 (rs8127225)	*TGF-βR1* (rs6478974)			0.02	0.24
TT	TT/T**A**	1699/2156	1.00 (REF)		
TT	**AA**	1185/1371	**1.14 (1.03-1.26)**		
T**C**/**CC**	TT/T**A**	417/434	**1.20 (1.03-1.39)**		
T**C**/**CC**	**AA**	222/241	1.17 (0.97-1.43)		
RUNX1 (rs7279123)	*TGF-β1* (rs4803455)			0.007	0.22
CC	CC	770/921	1.00 (REF)		
CC	C**A**/**AA**	1274/1545	0.93 (0.82-1.04)		
C**T**/**TT**	CC	417/477	1.02 (0.87-1.19)		
C**T**/**TT**	C**A**/**AA**	748/1006	**0.81 (0.71-0.93)**		
RUNX1 (rs7279123)	*TGF-βR1* (rs6478974)			0.10	0.66
CC	TT/T**A**	1842/2226	1.00 (REF)		
CC	**AA**	393/402	1.15 (0.99-1.34)		
C**T**/**TT**	TT/T**A**	1029/1291	0.94 (0.85-1.04)		
C**T**/**TT**	**AA**	243/275	1.02 (0.85-1.23)		

[a]Minor alleles (**bolded text**): rs4803455=A; rs6478974=A; rs7279383=G; rs2268288=C; rs8127225=C; rs7279123=T
[b]OR and 95% CI adjusted for age, study, and genetic admixture
[c]Wald-p for model; [d] Interaction p-value for interaction term in model (SNP*SNP)

Results for interactions between *RUNX2* (rs12209785, rs10948238, rs13201287)**TGF-β1* (rs4803455); and *RUNX2*TGF-βRI* (rs6478974) are found in Table 20. There were no significant interaction terms. There were significant associations among *RUNX2*TGF-β1* genotypes; the combination of 4 minor alleles was associated with an increase in risk for rs12209785* rs6478974 (**GG-AA**, OR=1.61 95% CI 1.08-2.40); rs10948238* rs6478974 (**TT-AA**, OR=1.42 95% CI 1.08-1.87); and rs13201287* rs6478974 (**AA-AA**, OR=1.48 95% CI 1.01-2.17), while there were no significant associations among the combinations *RUNX2*TGF-β1* genotypes.

In contrast to the other *RUNX* and *RUNX2* genes, there were two interactions for *RUNX3*TGF-βRI* found (Table 21). For rs7517302*rs6748974 (p=0.003), risk was similar for two categories, each with a differing homozygous minor allele group: TT/T**C**-**AA** and **CC**-TT/T**A** (OR=1.20-1.23). For rs906296*rs6748974 (p=0.02), risk was also increased among two categories, CC-**AA** and C**G**/**GG**-TT/T**A** (OR=1.20-1.26). While the other SNP-SNP interactions were not significant, an association was observed for when there is one homozygous minor allele group for rs4478762*rs6478974, GG/G**A**-**AA** and **AA**-TT/T**A**, OR=1.14 and 1.60, respectively.

Table 20. Interactions between RUNX2, TGF-β, TGF-βRI genes and breast cancer

Gene (SNP)[a]	Gene (SNP)[a]	Cases/Controls	Combined Risk OR (95% CI)[b]	p^c	*p-int*[d]
RUNX2 (rs12209785)	*TGF-β1* (rs4803455)			0.08	0.63
AA/A**G**	CC	1094/1298	1.00 (REF)		
AA/A**G**	C**A**/**AA**	1888/2390	0.90 (0.81-1.00)		
GG	CC	98/100	1.18 (0.89-1.59)		
GG	C**A**/**AA**	143/171	0.97 (0.77-1.23)		
RUNX2 (rs12209785)	*TGF- βR1* (rs6478974)			0.05	0.19
AA/A**G**	TT/T**A**	2681/3296	1.00 (REF)		
AA/A**G**	**AA**	582/634	1.10 (0.98-1.25)		
GG	TT/T**A**	201/233	1.08 (0.89-1.31)		
GG	**AA**	57/43	**1.61 (1.08-2.40)**		
RUNX2 (rs10948238)	*TGF-β1* (rs4803455)			0.03	0.77
CC/C**T**	CC	998/1198	1.00 (REF)		
CC/C**T**	C**A**/**AA**	1701/2180	0.90 (0.81-1.00)		
TT	CC	195/201	1.16 (0.94-1.44)		
TT	C**A**/**AA**	331/381	1.00 (0.84-1.19)		
RUNX2 (rs10948238)	*TGF- βR1* (rs6478974)			0.03	0.37
CC/C**T**	TT/T**A**	2431/3025	1.00 (REF)		
CC/C**T**	**AA**	519/576	1.10 (0.97-1.26)		
TT	TT/T**A**	450/503	1.12 (0.97-1.28)		
TT	**AA**	120/101	**1.42 (1.08-1.87)**		
RUNX2 (rs13201287)	*TGF-β1* (rs4803455)			0.04	0.69
GG/G**A**	CC	1083/1290	1.00 (REF)		
GG/G**A**	C**A**/**AA**	1865/2371	0.90 (0.81-1.00)		
AA	CC	110/110	1.22 (0.93-1.61)		
AA	C**A**/**AA**	167/190	1.02 (0.82-1.28)		
RUNX2 (rs13201287)	*TGF- βR1* (rs6478974)			0.05	0.55
GG/G**A**	TT/T**A**	2646/3270	1.00 (REF)		
GG/G**A**	**AA**	580/628	1.11 (0.98-1.26)		
AA	TT/T**A**	238/261	1.14 (0.95-1.38)		
AA	**AA**	59/49	**1.48 (1.01-2.17)**		

[a]Minor alleles (**bolded text**): rs4803455=A; rs6478974=A; rs12209785=G; rs10948238=T; rs13201287=A
[b]OR and 95% CI adjusted for age, study, and genetic admixture
[c]Wald-p for model; [d] Interaction p-value for interaction term in model (SNP*SNP)

Table 21. Interactions between RUNX3, TGF-β, TGF-βRI genes and breast cancer

Gene (SNP)[a]	Gene (SNP)[a]	Cases/Controls	Combined Risk OR (95% CI)[b]	p^c	*p-int*[d]
RUNX3 (rs7517302)	*TGF-β1* (rs4803455)			0.03	0.39
TT/T**C**	CC	1005/1189	1.00 (REF)		
TT/T**C**	C**A**/**AA**	1660/2153	0.88 (0.79-0.97)		
CC	CC	188/211	1.04 (0.84-1.29)		
CC	C**A**/**AA**	372/408	1.03 (0.87-1.21)		
RUNX3 (rs7517302)	*TGF- βR1* (rs6478974)			0.002	**0.003**
TT/T**C**	TT/T**A**	2377/3000	1.00 (REF)		
TT/T**C**	**AA**	538/541	**1.23 (1.08-1.40)**		
CC	TT/T**A**	505/528	**1.20 (1.05-1.37)**		
CC	**AA**	101/136	0.92 (0.71-1.20)		
RUNX3 (rs906296)	*TGF-β1* (rs4803455)			0.009	0.93
CC	CC	732/906	1.00 (REF)		
CC	C**A**/**AA**	1225/1624	0.90 (0.79-1.01)		
C**G**/**GG**	CC	461/494	1.14 (0.97-1.34)		
C**G**/**GG**	C**A**/**AA**	807/937	1.01 (0.88-1.16)		
RUNX3 (rs906296)	*TGF- βR1* (rs6478974)			0.0007	**0.028**
CC	TT/T**A**	1736/2283	1.00 (REF)		
CC	**AA**	406/417	**1.26 (1.08-1.46)**		
C**G**/**GG**	TT/T**A**	1146/1245	**1.20 (1.09-1.33)**		
C**G**/**GG**	**AA**	233/260	1.15 (0.95-1.39)		
RUNX3 (rs4478762)	*TGF-β1* (rs4803455)			0.04	0.65
GG/G**A**	CC	1174/1387	1.00 (REF)		
GG/G**A**	C**A**/**AA**	2000/2529	0.89 (0.81-0.99)		
AA	CC	19/13	1.68 (0.82-3.42)		
AA	C**A**/**AA**	32/31	1.22 (0.74-2.02)		
RUNX3 (rs4478762)	*TGF- βR1* (rs6478974)			0.03	0.23
GG/G**A**	TT/T**A**	2835/3491	1.00 (REF)		
GG/G**A**	**AA**	634/670	**1.14 (1.01-2.46)**		
AA	TT/T**A**	49/38	**1.60 (1.05-2.46)**		
AA	**AA**	5/7	0.86 (0.27-2.70)		

[a]Minor alleles: rs4803455=A; rs6478974=A; rs7517302=C; rs906296=G; rs4478762=A
[b]OR and 95% CI adjusted for age, study, and genetic admixture
[c]Wald-p for model; [d] Interaction p-value for interaction term in model (SNP*SNP)

There were no significant interaction terms when evaluating *ERα* (rs1801132 and rs3798577)**TGF-β* genes (genes/SNPs evaluated in previous interactions) (Table 22b, *See Appendix*). Most of the significant associations among combined genotypes were observed when homozygous variants for rs1801132 and rs3798577 were combined with the homozygous wild-type for other SNPs. These effects were expected, because when evaluated alone, these variant genotypes have similar magnitudes to the ones observed.

There were, however, notable differences among few combined genotypes (Table 22a). (Note: **bold** text indicates minor allele). Two significant associations had a >3-4-fold increase in risk for *ERα*TGF-β1* [rs1801132*rs4803455, **GG**-CC, OR=3.12) and *ERα*TGF-βRI* (rs1801132*rs6478974, **GG-AA**, OR=4.01)]. An increase in risk, although not as large, was also seen for *ERα*TGF-βRI* (rs3798577*rs6478974, **CC**-TT/T**A**, OR=1.45 95% CI 1.08-1.95). There was >2-fold increase in risk with the combination of homozygous variants for *ERα*RUNX1* [rs1801132*rs7279383, **GG**-C**G**/**GG**, OR=2.42) and *ERα*RUNX2* (rs3798577*rs12209785, **CC-GG**, OR=2.13). An increase in risk was also observed for *ERα*RUNX3* [(rs1801132*rs7517302, **GG-CC**, OR=3.37) and (rs3798577*rs906296, **CC**-C**G**/**GG**, OR=1.58)].

Table 22a. SNP-SNP interactions between ERα and TGF-β signaling genes (*Abbreviated table*)

			Combined Risk		
Gene (SNP)[a]	Gene (SNP)[a]	Cases/Controls	OR (95% CI)[b]	p^c	*p-int*[d]
ERα (rs1801132)	*TGF-β1* (rs4803455)			0.02	0.11
CC/C**G**	CC	202/198	1.00 (REF)		
CC/C**G**	C**A**/**AA**	428/474	0.89 (0.70-1.13)		
GG	CC	**22/7**	**3.12 (1.30-7.48)**		
GG	C**A**/**AA**	31/26	1.19 (0.68-2.08)		
ERα (rs1801132)	*TGF-βR1* (rs6478974)			0.06	0.21
CC/C**G**	TT/T**A**	497/537	1.00 (REF)		
CC/C**G**	**AA**	133/135	1.07 (0.82-1.40)		
GG	TT/T**A**	42/30	1.53 (0.94-2.49)		
GG	**AA**	**11/3**	**4.01 (1.11-14.47)**		
ERα (rs3798577)	*TGF-βR1* (rs6478974)			0.08	0.30
TT/T**C**	TT/T**A**	413/468	1.00 (REF)		
TT/T**C**	**AA**	113/109	1.18 (0.89-1.59)		
CC	TT/T**A**	**126/99**	**1.45 (1.08-1.95)**		
CC	**AA**	31/29	1.22 (0.72-2.07)		
ERα (rs1801132)	*RUNX1* (rs7279383)			0.06	0.18
CC/C**G**	CC	437/450	1.00 (REF)		
CC/C**G**	C**G**/**GG**	193/222	0.90 (0.71-1.24)		
GG	CC	32/24	1.38 (0.80-2.38)		
GG	C**G**/**GG**	**21/9**	**2.42 (1.10-5.36)**		
ERα (rs3798577)	*RUNX2* (rs12209785)			0.06	0.77
TT/T**C**	AA/A**G**	492/549	1.00 (REF)		
TT/T**C**	**GG**	34/28	1.35 (0.81-2.27)		
CC	AA/A**G**	**144/121**	**1.34 (1.02-1.75)**		
CC	**GG**	**13/7**	**2.13 (0.84-5.40)**		
ERα (rs1801132)	*RUNX3* (rs7517302)			0.05	0.34
CC/C**G**	TT/T**C**	502/552	1.00 (REF)		
CC/C**G**	**CC**	128/120	1.17 (0.89-1.54)		
GG	TT/T**C**	41/29	1.56 (0.95-2.54)		
GG	**CC**	**12/4**	**3.37 (1.08-10.53)**		
ERα (rs3798577)	*RUNX3* (rs906296)			0.03	0.65
TT/T**C**	CC	299/359	1.00 (REF)		
TT/T**C**	C**G**/**GG**	227/218	1.25 (0.98-1.60)		
CC	CC	**93/79**	**1.43 (1.02-2.00)**		
CC	C**G**/**GG**	**64/49**	**1.58 (1.05-2.37)**		

[a]Minor alleles: rs1801132=G; rs3798577=C; rs4803455=A; rs6478974=A; rs12209785=G; rs7517302=C; rs906296=G; rs7279383=G

[b]OR and 95% CI adjusted for age and genetic admixture

[c]Wald-p for model; [d] Interaction p-value for interaction term in model (SNP*SNP)

Cumulative Effect of Risk Alleles and Association with Breast Cancer

A GRS was created for those significant ($p<0.05$) or marginally significant ($p<0.15$) SNPs was split into those that were associated with increased risk and reduced risk of breast cancer. The number of risk alleles (0, 1 or 2) was summed across SNPs. Table 23 shows GRS-1, creating by summing risk alleles for TGF-B signaling SNPs associated with reduced risk (n=4). GRS-1 range was 0-7 and the per-allele effect was significantly associated with a reduced risk (OR=0.92 95%CI 0.88-0.96). When evaluated as a categorical variable, compared to those individuals with ≤1 risk allele, those with ≥5 risk alleles had the largest inverse association (OR=0.65 95% CI 0.49-0.87).

Table 23. Genetic Risk Score 1: TGF-β signaling SNPs associated with reduced risk

Gene	SNP	Risk allele[a]	OR (95% CI)-per allele[b]	*p-trend*[c]
RUNX1	rs7279383	G	0.91 (0.83-1.00)	0.05
RUNX1	rs7279123	T	0.94 (0.87-1.02)	0.13
TGF-β1	rs4803455	A	0.95 (0.89-1.02)	0.15
TGF-βR1	rs10733710	A	0.94 (0.88-1.01)	0.08

[a] Allele associated with breast cancer when compared to wild-type allele (referent)
[b] OR (95% CI) is SNP entered as continuous; adjusted for age, study, genetic admixture
[c] Wald-p for SNP entered as continuous, per-allele effect

Genetic Risk Score (GRS)-1

	GRS-1[d]	Cases/Controls	OR (95% CI)[e]	*p-value*[f]
Categories*	≤1	1114/1287	1.00 (Referent)	
	2	1048/1235	0.96 (0.86-1.08)	0.51
	3	669/904	0.83 (0.73-0.95)	0.005
	4	296/384	0.86 (0.72-1.02)	0.08
	≥5	81/138	0.65 (0.49-0.87)	0.003
Trend[g]	0-7 Alleles	3208/3948	0.92 (0.88-0.96)	0.0003

[d] Number of risk allele categories are based on control population
[e] OR (95% CI) adjusted for age, study, genetic admixture
[f] Wald-p for each GRS-1 category; [g] Wald-p for GRS-1 entered as continuous, per-allele effect

Details regarding the SNPs summed to create GRS-2 are shown in Table 24. There was one SNP from *RUNX1, RUNX2, RUNX3* and *TGF-βR1* used to create GRS-2, all associated with increased risk. GRS-2 range was 0-7 risk alleles and the per-allele effect was significantly associated with a slight increase in risk (OR=1.08 95% CI 1.04-

1.12). Compared to those individuals with ≤1 risk allele, those with 2 or 3 risk alleles had the similar risk (OR=1.17 and 1.19, respectively). Individuals with 4 or ≥5 risk alleles also had similar risk, which was larger in magnitude than the other categories (OR=1.34).

Table 24: Genetic Risk Score 2: TGF-β signaling SNPs associated with increased risk

Gene	SNP	Risk allele[a]	OR (95% CI)-per allele[b]	*p-trend*[c]
RUNX1	rs8127225	C	1.08 (1.00-1.17)	0.05
RUNX2	rs10948238	T	1.05 (0.99-1.13)	0.11
RUNX3	rs906296	G	1.11 (1.03-1.20)	0.009
TGF-βR1	rs6478974	A	1.07 (1.00-1.14)	0.04

[a] Allele associated with breast cancer when compared to wild-type allele (referent)
[b] OR (95% CI) is SNP entered as continuous; adjusted for age, study, genetic admixture
[c] Wald-p for SNP entered as continuous, per-allele effect

Genetic Risk Score (GRS)-2

	GRS-2[d]	Cases/Controls	OR (95% CI)[e]	*p-value*[f]
Categories*	≤1	738/1027	1.00 (Referent)	
	2	1069/1269	1.17 (1.03-1.33)	0.014
	3	940/1102	1.19 (1.05-1.35)	0.008
	4	560/580	1.34 (1.15-1.55)	0.0002
	≥5	211/218	1.34 (1.08-1.66)	0.007
Trend	0-7 Alleles	3518/4196	1.08 (1.04-1.12)	<0.0001

[d] Number of risk allele categories are based on control population
[e] OR (95% CI) adjusted for age, study, genetic admixture
[f] Wald-p for each GRS-2 category; [g] Wald-p for GRS-2 entered as continuous, per-allele effect

Lastly, GRS-3 was created to evaluate the cumulative effect between pathways using the four SNPs from GRS-2 and summing with number risk alleles from *ERα* (rs1801132 and rs6913578) (Table 25). GRS-3 range was 0-10 risk alleles and the per-allele effect was also significantly associated with a slightly higher risk (OR=1.14 95% CI 1.04-1.25). Compared to those individuals with ≤3 risk alleles, those with 5 or 6 had the highest risk (OR=1.67 and 1.52, respectively). Those with ≥7 risk alleles only had a slight increase in risk (OR=1.11), which was not significant.

Table 25: Genetic Risk Score 3: ERα and TGF-B signaling SNPs associated with increased risk

Gene	SNP	Risk allele[a]	OR (95% CI)[b]	*p-trend*[c]
RUNX1	rs8127225	C	1.08 (1.00-1.17)	0.05
RUNX2	rs10948238	T	1.05 (0.99-1.13)	0.11
RUNX3	rs906296	G	1.11 (1.03-1.20)	0.009
TGF-BR1	rs6478974	A	1.07 (1.00-1.14)	0.04
ERα	rs1801132	G	1.18 (0.99-1.40)	0.07
ERα	rs6913578	C	1.14 (0.97-1.34)	0.11

[a] Allele associated with breast cancer when compared to wild-type allele (referent)
[b] OR (95% CI) is SNP entered as continuous; adjusted for age, study, genetic admixture
[c] Wald-p for SNP entered as continuous, per-allele effect

Genetic Risk Score (GRS)-3

	GRS-2[d]	Cases/Controls	OR (95% CI)[e]	*p-value*[f]
Categories*	≤3	314/360	1.00 (Referent)	
	4	162/190	0.99 (0.76-1.28)	0.93
	5	120/83	1.67 (1.21-2.29)	0.002
	6	66/50	1.52 (1.02-2.26)	0.04
	≥7	21/22	1.11 (0.60-2.06)	0.74
Trend	0-10 Alleles	683/705	1.14 (1.04-1.25)	0.007

[d] Number of risk allele categories are based on control population
[e] OR (95% CI) adjusted for age and genetic admixture
[f] Wald-p for each GRS-3 category; [g] Wald-p for GRS-3 entered as continuous, per-allele effect

GRS-3 was further evaluated to test for associations with breast cancer defined by ER status (Table 26). Risk was similar for ER+ and ER- tumors with the exception of those individuals with 6 risk alleles: there was a significantly higher risk for ER+ breast cancer (OR=1.61) but not ER- (OR=1.17). For those with ≥7 risk alleles there was a higher risk for ER- breast cancer (OR=1.44), but not ER+ (OR=1.13); however, these estimates did not reach statistical significance.

Table 26. Genetic Risk Score (GRS)-3 and breast cancer defined by ER status

GRS-3[a]	Controls[b]	ER+ (N)	ER+ [c]	ER- (N)	ER- [c]
≤3	360	159	1.00 (Referent)	45	1.00 (Referent)
4	190	77	0.92 (0.67-1.27)	20	0.83 (0.47-1.45)
5	83	53	1.45 (0.98-2.15)	18	1.68 (0.92-3.06)
6	50	36	1.61 (1.01-2.57)	7	1.17 (0.50-2.76)
≥7	22	11	1.13 (0.54-2.40)	4	1.44 (0.47-4.39)
0-10(trend)[c]	705	336	1.11 (1.02-1.21)	94	1.07 (0.93-1.24)

[a] Number of risk allele categories are based on control population
[b] ER were compared with 705 controls; [c] OR (95% CI) adjusted for age and genetic admixture

DISCUSSION

The primary aim for this dissertation study was to evaluate genetic variation in *TGF-β* signaling genes (*TGF-β1, TGF-βRI, RUNX1, RUNX2,* and *RUNX3*) (n=7,733), *ERα* gene (n=1,409) and breast cancer risk among Hispanic and NHW women who participated in one of three population-based case-control studies: 4-CBCS, MBCS, or SFBCS. Ancestry informative markers were also evaluated to make a distinction between proportion of European and Native American ancestry (genetic admixture strata), which serves as an indicator of ethnicity. Due to the hypothesized cross-talk within and across the *TGF-β* signaling pathway and ERα, there was particular interest to evaluate the association with breast cancer defined by ER status.

Overall Associations with Breast Cancer

Summary Results

After adjustment for age, study, and genetic admixture, nine SNPs [*RUNX1* (rs7279383 and rs8127225); *RUNX2* (rs10948238 and rs13201287); *RUNX3* (rs906296); *TGF-β1* (rs4803455); *TGF-βR1* (rs6478974), and *ERα* (rs1801132 and rs3798577)] were associated with overall breast cancer risk. However; only two SNPs remained significantly associated after adjustment for multiple comparisons, one with a slight inverse association, *TGF-β1* (rs4803455), while the other slightly increased risk, *RUNX3* (rs906296).

Previous Literature

Genes from the *TGF-β* super-family play important roles in regulating cellular processes such as proliferation, differentiation, adhesion, migration, and survival [155]. This pathway has been found to be the most commonly altered cellular signaling pathway in cancer, , and may suppress or promote tumors depending on the inactivity of core components, making it an attractive candidate when studying cancer etiology [93, 155]. Several epidemiological case-control studies have investigated associations between select SNPs in *TGF-β1, TGF-βR1* and breast cancer risk [96, 101-102, 156-167]. Data are inconsistent and most studies are underpowered and include NHW or Asian women only. There have been several meta-analyses conducted for *TGF-β1* and *TGF-βR1* SNPs.

The most common *TGF-βR1* SNP evaluated is *6A or rs1466445, which results from the deletion of three alanines within a nine-alanine (*9A) stretch in exon 1 [168]. This variant has been associated with decreased expression of *TGF-βRI* [169]. TGF-βR1*6A was found to increase breast cancer risk (per-allele OR=1.15) in a comprehensive meta-analysis conducted in 2012 including 17 case-control studies [170], including NHW, Asian or Indian populations and lacked inclusion of Hispanic women. The Nurses' Health Study was the largest with ~1,200 cases, predominately NHW [101]. Zhang and colleagues estimated that, given the high carrier rate (general population frequency ~14% heterozygote ; 0.5% homozygous variant), the population attributable risk (PAR) was 4.9% (2.7%-7.2%) for all breast cancers [164]. Although this variant was not genotyped for the present study, 5 tagSNPs were evaluated. *TGF-βR1* (rs6478974) was associated with a slight increase in risk (OR_{AA}=1.13 95% CI 1.00-1.28, p=0.05), although significance did not remain after adjustment for multiple comparisons. Two

studies have found that *TGF-βR1* (rs11568785) is in strong LD with *TGF-βR1**6A and suggest that it may mediate the functionality of *TGF-βR1**6A [169, 171]. In the present study *TGF-βR1* (rs11568785) was associated with a modest increase in risk (OR=1.41 95% CI (0.74-2.69). Although it was not statistically significant, theoretically it could be associated with a non-synonymous SNP (*A). However, it is beyond the means of this study to test for LD between the two SNPs.

The most widely reported *TGF-β1* SNPs are rs1982073 (T869C), located on exon 1, which results in a leucine to proline substitution at codon 10(Leu10Pro), and rs1800469 (C-509T) in the promoter region [10]. The variant allele of rs1982073 (C) has been found to be associated with increased serum and plasma concentrations of *TGF-β1.* It has been hypothesized that women carrying this allele may be at a lower risk of breast cancer [159, 172-173]. However, two recent meta-analyses suggest that the per-allele effect of *TGF-β1* (rs1982073) is associated with a ~5-8% increased risk of breast cancer in NHW women [88, 103]. Wei and colleagues [174] also conducted a meta-analysis on *TGF-β1* (rs1982073) and reported a null association (OR=1.02 95% CI 0.94-1.07); however they did not include one large study with data from ~13,000 cases in the Breast Cancer Association Consortium (BCAC) that found a significant association with this SNP and breast cancer (per allele OR=1.08 95% CI 1.04-1.31) [156]. Le Marchand, et al. [157] was the only study to report an estimate for Hispanic women in the Multiethnic Cohort study, including post-menopausal women only. Compared to Hispanic women with the TT genotype, women with CC-genotype had reduced risk (OR=0.81 95% CI 0.52-1.27). However, the number of women was small (cases=67, controls=179) and the analysis was underpowered.

Three meta-analyses were conducted for *TGF-β1* (rs1800469) and breast cancer risk in 2010 [103, 175-176], although there were differences in studies used for a pooled estimate, results were similar. Niu, et al. [176] included 9 study populations [101, 159-162, 177], Qi, et al. [103] included 8 of the 9 [101, 159-162, 177], and Woo, et al. [175] included 7 of 9 [101, 159-160, 162]. Woo, et al. did not include two studies [161, 177], which contributed ~10,700 women to the other two pooled estimates. It is important to note here that only one of the NHW populations included in this meta-analysis was from the US [101], while others studies represented NHW populations in Finland, Germany, Poland, and United Kingdom [103, 175-176]. Evaluating the recessive model (TT *vs.* CC/CT), the *TGF-β1* (rs1800469) was found to have no effect in all three meta-analyses [(OR_{CC}=1.00 (95% CI 0.89-1.15) [176]; (OR_{CC}=1.00 (95% CI 0.88-1.14) [103]; and (OR_{CC}=0.92 (95% CI 0.83-1.03) [175]. Data from the present study suggest there is a moderate LD (r^2=0.67) between *TGF-β1* (rs1800469) and *TGF-β1* (rs4803455) in both Hispanic and NHW women. After adjustments for multiple comparisons, a significant inverse association of breast cancer was observed for *TGF-β1* (rs4803455) ($OR_{CA/AA}$ =0.89) and a non-significant positive association for *TGF-β1* (rs1800469) (OR_{TT} =1.08). When evaluated by Native American ancestry, there was a positive association for *TGF-β1* (rs1800469) among women with 29-71% ancestry (OR_{TT}=1.29), while the association was inverse for *TGF-β1* (rs4803455) ($OR_{CA/AA}$=0.86), which did not remain significant after adjustment for multiple comparisons. In contrast to findings from this study, Scollen and colleagues found that *TGF-β1* (rs4803455) increased risk in a co-dominant model (OR_{AA}=1.21 95%CI 1.02-1.43), in a population including ~4,500 NHW women [88].

TGF-β1 (rs1982073 and rs1800469) have also been linked to a higher expression of *TGF-β1* in breast tumors [102, 159]. The Asian population of the Shanghai Breast Cancer study [160] was included in the three meta-analyses and reported that rs1982073 and rs1800469 were in high LD. They evaluated the association of breast cancer defined by stage of diagnosis. Compared to the homozygous wild-types, carries of homozygous wild-types were inversely associated with early stage (0/I) of disease and there was a non-significant increase for later stage (II-IV) for both SNPs.

All three *RUNX* genes have been shown to be important in colorectal cancer [178]. Little is known about the role of genetic variation in these genes in breast cancer etiology, although research is warranted given their role in the signaling pathway mediated by *TGF-β1*. In the present admixed population, before multiple comparison evaluation, one SNP was associated with reduced risk: *RUNX1* (rs7279383); while four SNPs were significantly associated with increased risk: *RUNX1* (rs8127225), *RUNX2* (rs10948238 and rs13201287), and *RUNX3* (rs906296). These findings appear to be the first for an association between genetic variation in the *RUNX* genes and breast cancer risk, providing additional support for the *TGF-β* signaling pathway and breast cancer etiology.

The gene *ERα* regulates the biological function of the steroid hormone estrogen and is an important predictive and prognostic factor in breast cancer [7]. Most published association studies have evaluated the most common SNPs of ERα: PvuII (397T > C, rs2234693) and the XbaI (351A > G, rs9340799) restriction fragment length polymorphisms, both located in intron 1 and found to be in LD [179]. There are mixed results for these two SNPs reported from several studies [180-187]. In a meta-analysis of

11 case–control studies (cases=8,255; controls=13,164) there was a slight non-significant decrease in risk observed for PvuII (rs2234693) with the TT *vs.* CC genotype (OR=0.92 95% CI 0.86-0.99). In a meta-analysis of 10 case-control studies (cases=8,645; controls=12,805) there was a null association (OR=0.99 95% CI 0.90-1.08) for GG *vs.* AA genotype of XbaI (rs9340799) [188]. In one of the studies included in this dissertation, 4-CBCS, Slattery and colleagues evaluated XbaI (rs9340799) and did not find an association with breast cancer in Hispanic or NHW women overall [187].

The present study evaluated five *ERα* SNPs; two of which (rs1801132 and rs3798577) were found to be associated with a modest increase in breast cancer before adjustment for multiple comparisons. *ERα* (rs1801132) is a well-studied, synonymous SNP in codon 325 of exon 4 (hormone binding domain), and has been evaluated in several previous population-based studies [180, 182-183, 185, 189-193]. Li and colleagues [188] conducted a meta-analysis including seven of these studies [180, 182-183, 185, 190, 192-193] (cases=5,649, controls=6,856) and reported a non-significant reduced risk for the dominant model ($OR_{CG/GG}$=0.92 95% CI 0.95-0.99), however they were not able to pool estimates for a recessive model. Since this meta-analysis, Anghel and colleagues [189] did not find an association with rs1801132 and overall risk, however there was a significant association for age at diagnosis (diagnosis >50 years, p=0.02). There was a non-significant increased risk for rs1801132 in a recent study in a Chinese population ($OR_{CG/GG}$=1.12 95% CI 0.90–1.40), however it was not significant [191]. The present study found a non-significant increase in the dominant model ($OR_{CG/GG}$=1.13 95% CI 0.91-1.39); although when evaluating the recessive model there was a higher and more significant association (OR_{GG}=1.72 95% CI 1.10-2.69), which is

in contrast to the meta-analysis. This dissertation study also found a significant increase in risk for ERα (rs3798577, OR_{CC}=1.36), which is located in the 3'-untranslated region (3'-UTR; a region of messenger RNA (mRNA)) of ERα. Although its functionality is not well established, this region seems to alter the *ERα* expression [189]. A meta-analysis, conducted by Li and colleagues [188], pooled data from three studies [180, 182-183] and no association was observed (estimate not given). Lastly, two studies have reported an increase in risk: a small study (n=192) conducted in Romania [189] reported a >2-fold increase in risk (per-allele); and a study conducted in the United Kingdom with ~3,900 women (per-allele OR=1.11) [194].

The present study did not find a significant association with breast cancer for *ERα* rs2046210, rs851984, or rs6913578. In contrast, these SNPs were reported to be associated with breast cancer in previous studies that included NHW, Japanese and/or Chinese populations. *ERα* (rs2046210) is located on 6q25.1 (1,440 base pairs upstream of the coding region of *ERα*). Zheng, et al. found a significant increase in risk in a GWAS of Chinese women (OR_{AA}=1.59 95% CI 1.40–1.81) as well as European women from the Nashville Breast Health Study (OR_{AA}=1.35 95% CI 1.06–1.71) [195]. Cai and colleagues pooled data from fourteen studies (cases=17,188, controls=14,660) and replicated the results of Zheng [196] in women of Chinese (OR_{AA}=1.64), Japanese (OR_{AA}=1.37) and European ancestry (OR_{AA}=1.18); although there was no association in African American women. Han et al. reported a significant increased association for the dominant model (GA/AA *vs.* GG, OR=1.47 95% CI 1.27–1.69). This study also supports an increase in risk with the dominant model ($OR_{GA/AA}$=1.22 95% CI 0.98-1.43), although of borderline significance. *ERα* (rs6913578) is 1,440 bp downstream of rs2046210. Cai

and colleagues suggested that *ERα* (rs6913578) may be a functional variant as they found that the minor allele (C) of rs6913578 significantly altered DNA-protein complex (II) intensity in both Human Embryonic Kidney 293 cells (HEK293) and Michigan Cancer Foundation-7 cells (MCF-7, breast cancer cell line). They also report rs6913578 to be in high LD with rs2046210 in Chinese (r^2=0.91) and European-ancestry (r^2=0.83), and found an increased risk in both populations, OR_{CC}=1.54 and 1.31, respectively [196]. Again, data from the present study appear to support an increase in risk (OR_{CC}=1.26 95% CI 0.88-1.81), although not statistically significant.

Lastly, *ERα* (rs851984), located near the promoter region [183], was evaluated in the present study and found to have a slight, non-significant positive association (OR_{AA}=1.16 95% CI 0.84-1.59). In contrast, MARIE-GENICA (a pooled analysis of postmenopausal women (cases=3,149, controls=5,489) from two population-based studies in Germany found a significant per-allele effect (OR_A=1.13 95% CI (1.03-1.25) [197].

Biological Mechanism

Breast Cancer Initiation

It is widely acknowledged that components of *TGF-β* signaling pathway, *ERα* and their target genes play a role in breast development and can either support or inhibit the growth of tumor cells. Understanding the complex functioning of these genes and the relationship with breast cancer is a significant issue for a further evaluation of predictive risk factors.

In most tumor cells, genetic variation in key members of the pathway can cause resistance to the growth inhibitory effects of *TGF-β* signaling [96-97]. Exact

mechanisms for resistance remains unknown, although researchers have hypothesized, through evidence in gastric, pancreatic, and colon cancer studies, that there may be decreased expression of its receptors (*TGF-βRI* and *TGF-βRII*) on the cell surface or increased expression of inhibitory SMADs (*I-SMAD6* or *I-SMAD-7*) in the extracellular matrix, inhibiting the signaling function [87]. Some researchers suggest that reduced expression or inactivation of *TGF-β* signaling could be caused by oncoproteins such as *p53* [98] or decreased expression of other tumor suppressors that regulate the pathway such as *RUNX3* [99].

The *RUNX* genes are established down-stream effectors of *TGF-β* signaling. *RUNX3* is best known for its role as a tumor suppressor in the gastrointestinal tract. More recently, research has found it to be a tumor suppressor in breast cancer, as *RUNX3* mRNA is consistently under expressed in tumor cells compared to normal cells [136]. This could be a result of protein mislocalization, which could potentially cause a disturbance in the mechanism that controls nuclear transport of *RUNX3* resulting in the atypical localization of RUNX3in the cytoplasm. Epigenetic changes, most commonly, hypermethylation of the CpG island in the promoter region of *RUNX3* can cause silencing or functional inactivation of the tumor suppressor function [124-125].

RUNX1 is established as a tumor suppressor in hematopoietic malignancies. It has been reported to be expressed in luminal and basal epithelial cells in normal breast tissue, but to be deficient in breast tumor tissue [107, 114]. Although its role in breast tissue is understudied, research suggests *RUNX1* may be deregulated in breast carcinogenesis. Janes and colleagues report that *RUNX1* and a family of transcription factors, called *FOXO*, have an antagonistic relationship. Specifically, when cells are

RUNX1-deficient, they are under oxidative stress and FOXO expression increases to stabilize and support cell proliferation [107].

In contrast to components of the *TGF-β*, *ERα* supports cell proliferation. When estrogen binds to nuclear *ERα*, the estrogen-*ERα* complex then binds to sequences known as estrogen response elements (ERE) and effects transcriptional activity [78]. The role *ERα* plays in breast carcinogenesis is well studied and two biological mechanisms have been described [7]. First, binding of estrogen to the ERα stimulates mammary cell proliferation, increasing in cell division and DNA replication leading to spontaneous errors or mutations. As an example in ER+ cells, estrogen down-regulates E-cadherin, a mediator of cell-cell interactions that plays an important role in tumor suppression in the breast. Lowered E-cadherin expression in both normal and tumor epithelial cells in the breast is caused via a decrease in promoter activity and subsequent mRNA levels [198].

The second mechanism involves metabolism of endogenous estrogen producing genotoxic by-products that could directly harm DNA [199]. As an example, ring hydroxylation at the C-2 and C-4 positions of endogenous estrogens forms catechol estrogens and then reactive quinone metabolites that have been increasingly associated with estrogen-induced breast cancer that cause oxidative damage to DNA [200]. This is further supported by an epidemiologic study that reported postmenopausal women with low-activity genotype for catechol-*O*-methyltransferase (*COMT*, gene responsible for inactivating catechol estrogens) are at increased risk of breast cancer compared to those with the highly active genotype [201].

There are also hormone-independent tumors (*ERα* negative) that can develop, whose mechanism is not understood. Suggested mechanisms include loss of ligand-

binding and responsiveness of estrogen via genetic mutations in ERα. Variants at the mRNA level with alternative splicing can yield deletion of exon 3, 5, or 7 causing repression of *ERα* protein. Amino acids in the AF-1 domain of ERα are phosphorylated by activation of a signaling cascade downstream of receptor tyrosine kinases. The phosphorylated ERα is then able to regulate the transcription of target genes via protein-protein interaction with other transcription factors or corregulators [202]. In addition, DNA methylation of *ERα* may control its expression and therefore play an important role in the hormone-independent breast cancer [203]. Ultimately, the result of these mechanisms disturbs the normal functions of the cell cycle, apoptosis and DNA repair, leading to breast cancer development and promotion.

Breast Cancer Promotion

There is also evidence that these genes can influence breast cancer progression. When *TGF-β1*and *TGF-βRI* are overexpressed following tumor initiation, they promote metastasis [87]. In order for a tumor to metastasize, tumor cells must have the ability to migrate in and out of blood vessels and invade surrounding tissues. This has also been called the epithelial to mesenchymal differentiation of tumors and *TGF-B* protein is able to induce this transition in cultures of breast epithelial cells [204]. This concept of 're-programming' the *TGF-B* protein has been suggested to be in response to a disturbance by epigenetic mechanisms or genetic alterations, such as the activation of *Ras*, a subfamily of *GTPase* proteins that transmit signals within cells [204]. The combined effect of *Ras/MAP* kinase signaling can induce expression of *TGF-B1*, enhancing the signaling pathway, which could explain the increased levels of *TGF-B1* in breast tumors and subsequent tumor invasion [205-206].

The formation of a network of blood vessels (angiogenesis) is necessary to provide nutrients to the tumor cells. High levels of *TGF-β1* mRNA have been found to be associated with increased microvessel density in human breast tumors [207]. In this environment *TGF-β1* has been found to induce the expression of an angiogenic growth factor: vascular endothelial growth factor (*VEGF*), which acts to stimulate proliferation in endothelial cells [100, 208]. Research suggests an indirect role for *RUNX2* and breast metastasis through alteration of *VEGF* as well [120]. Conversely, *RUNX3* is found to be down-regulated in most breast cancers, Chen and colleagues reported that when *RUNX3* is activated or overexpressed the invasive potential of breast cancer cells (MDA-MB-231) is reduced in-vitro [209].

The role of *ERα* signaling and breast cancer promotion is not well understood. A potential mechanism may involve the recruitment of co-factors by *ERα* that have a negative effect on cell motility and invasion, although these results have been inconsistent are dependent on complex interactions with other signaling pathways [210]. Disruptions in the *ERα*-signaling pathway could lead to estrogen-dependent or estrogen-independent mechanisms involving loss of hormone responsiveness, reduction of tumor suppressor functions, interaction with growth factors (i.e. *VEGF*), and nuclear proto-oncogenes (normal genes mutated to be oncogenes, i.e. *c-fos* and *c-myc*), to name a few [200, 203, 211].

Stratification by Proportion Native American Ancestry

Summary Results

Unlike previous studies evaluating *TGF-β1* and *TGF- βRI* and breast cancer risk, this dissertation study was able to test for interaction effects with genetic admixture.

Unique associations were observed by genetic admixture. Predominantly, *TGF-β* signaling genes and *ERα* did not differ by genetic admixture; there were however, three SNPs were more associated with increased breast cancer risk in women of moderate (*RUNX3* (rs906296) and *TGF-β1* (1800469)) and high (*RUNX1* (rs7279383)) Native American ancestry after adjustment for multiple comparisons. There was also a significant interaction found between *RUNX1* (rs7279383) and genetic admixture (p=0.04).

Previous Literature

Previously, Slattery and colleagues [8] reported that women with higher NA ancestry were at a reduced risk compared to women with a more European ancestry. In the present study, the difference in risk between selected SNPs and admixture strata may support the hypothesis that genetic variation related to important factors such as cell growth and hormones can influence breast cancer differently in Hispanic and NHW women.

Published data from the 4-CBCS also indicate that there are ethnic differences for associations of select SNPs and breast cancer: the interleukin-6 (IL6) SNPs had a greater association with risk among Hispanic than NHW women [212]; a higher risk among NHW was observed with serum Insulin-like growth factor 1 (IGF-1) compared to Hispanic women [213]; the beta-2-adrenergic receptor (ADRB2) haplotype was associated with a greater risk in NHW than Hispanic women having a BMI ≥25 kg/m2 [214]. Other data from the BCHD study also indicate bone morphogenetic proteins (BMP1, BMP6, BMPR1B, BMPR2), which are members of the TGF-β signaling pathway, differ across genetic admixture groups [215]. One plausible explanation for

these differences in risk between ethnic groups could be that there is a region including unmeasured biologically functional variants that differs by ethnicity that contributes to breast cancer susceptibility.

Stratification by Menopausal status

Summary Results

In general, breast cancer risk did not differ by menopausal status in the present study. There were no significant interactions with menopausal status. Within menopausal strata the The CG/GG genotype of *RUNX3* (rs906296) was significantly associated greater risk among pre-menopausal (OR=1.33), but not post-menopausal women. Although *ERα* (rs1801132) was not associated with overall risk after adjustment for multiple comparisons, the risk for the GG-genotype was >2-fold among post-menopausal women, while risk for pre-menopausal women was attenuated and not significant.

Associations with Breast cancer Defined by ER/PR status

Summary Results

Data from the present study supports the theory that these genes are associated with breast cancer phenotypes as defined by their hormone receptors. SNPs in these genes were associated with specific breast cancer tumor phenotypes: *RUNX3* (2 SNPs) was associated with ER+/PR+ tumors (OR 1.18 and 1.90); *RUNX1* (2 SNPs) and *TGF-βR1* (1 SNP) were associated with ER+/PR- tumors (OR between 0.44 and 3.55); *RUNX3* (3 SNPs) was associated with ER-/PR+ tumors (OR between 2.52 and 2.88); and *RUNX1* (1 SNP), RUNX2 (5 SNPs), RUNX3 (2 SNPs) were associated with ER-/PR- tumors (OR between 0.77 and 2.31). After adjustment for multiple comparisons, four *RUNX*

SNPs remained significantly associated with an increased risk of ER-/PR- (n=3) and ER-/PR+ (n=1) tumors. Further evaluation of ER+/ER- tumors by menopausal status showed that risk for ER- tumors is significantly associated with >2-fold increase in post-menopausal women (*RUNX1*, rs2268288 and *RUNX2*, rs12333172). When determining whether risk of ER+/ER- breast differed by proportion Native American ancestry, there were a larger number of significant positive [ER+ (n=2), ER- (n=1)] and inverse associations [ER+ (n=1) ER- (n=1)] for those with moderate to high Native American ancestry, which were not observed in the low Native American ancestry group.

This study is the first to evaluate associations *TGF-β* and *RUNX* genes and risk of breast cancer defined by ER/PR tumor phenotypes. Although the present analysis included 1,962 cases with available data on ER/PR status, there was a lack of tumor phenotype data for the MBCS, limiting power when evaluating some tumor phenotypes, specifically ER-/PR+ (n=45 cases). However, these results strengthen the importance of these genes and estrogen-related associations with breast cancer.

Biological Mechanism: Estrogen Links Signaling Pathways

Genes in the TGF- *β* signaling pathway have been associated with estrogen via expression of estrogen receptors (ERα) and estrogen signaling pathways [133, 136, 216]. Although ERα and TGF- *β* have an opposing regulatory effect on normal cell proliferation (promotion and inhibition, respectively), a potential relationship in breast carcinogenesis has been elucidated. Several studies have provided evidence that receptor regulated SMADs (*R-SMAD2, R-SMAD3*) and common mediator (*Co-SMAD4*) comes into direct physical contact with *ERα* [129-132]. *Co-SMAD4* is found to be a mediator of crosstalk between *TGF-β* and *ERα* where it acts as a co-repressor of the transcription of

ERα, inhibiting tumor growth. Interestingly, *Co-SMAD4* has been found to induce apoptosis in *ERα*+ but not *ERα*– cells [133]. In the absence of *Co-SMAD4*, *ERα*-estrogen cell proliferation is enhanced [131]. Estrogen will act acting directly on the *TGF-β* signaling pathway to block the phosphorylation of *R-SMAD* 2/3 complex via ubiquitin-proteasome pathway [130]. Bierie and colleagues compared breast cancer expression signatures to *TGF-β* response signatures and found two associations: first, the *TGF-β* response signature was associated with *ERα*-negative tumors and poor prognosis; second, the absence of *TGF-β* response signature was found to be higher in *ERα*-positive tumors and was associated with a poor prognosis [135].

The *RUNX* transcription factors were reported to interact with estrogen signaling. *RUNX1* has been called an 'accessibility factor' for *ERα* binding sites, and may function as a regulator of *ERα* gene expression specifically in *ERα*-positive cells [216]. *RUNX1* has been found to suppress the oncogenic effects of estrogens mostly through this physical interaction with *ERα*, called 'tethering' [217]. *RUNX2* has been observed to interact with *ERα* involving two mechanisms: first, *RUNX2* decreases *ERα* mRNA and protein levels in breast cancer cells; and second, the interaction of *RUNX2* and *ERα*-binding domain results in decrease association of *ERα* with its target genes [218]. Using mouse models, Huang and colleagues showed that *RUNX3* may target *ERα* to function as a tumor suppressor by destabilizing the gene and inhibiting its expression [136]. An inverse relationship was observed between expressions of the two genes; a higher *RUNX3* expression was associated with lower *ERα* in *ERα*-positive cells and *vice versa*, while Lau and colleagues found the same effect in *ERα*-negative cells [125, 136].

Evaluation of Potential Crosstalk between Signaling Pathways

Summary Results

Evaluation of SNP-SNP interactions provided support that the multiplicative effect of SNPs in the *TGF-β* signaling pathway alters breast cancer susceptibility. Specifically there were two significant interactions between *RUNX3* and *TGF-βRI*. Two GRS evaluated the cumulative effect of risk alleles for SNPs that were found to have an inverse association and positive association with breast cancer. Findings for the GRS including four SNPs showed that the per-allele trend of risk alleles reduced risk (range=0-7 alleles) (OR=0.92 95% CI 0.88-0.96). There was also a higher risk observed for the GRS including 6 SNPs (range=0-10 alleles) (OR=1.14 95% CI 1.04-1.25), and was also positively associated with ER+ tumors but not ER- tumors. It is important to note that the p-values for these findings are biased due to the fact that the risk alleles for selected SNPs were based solely on this study population. To further test this hypothesis, a refinement of statistical methods would be necessary, such as permutation testing, allowing for the correct distribution of a test statistic under a null hypothesis, resulting in a true p-value [219]. Nevertheless, this is the first population-based observational study evaluating the crosstalk between *TGF-β* and *ERα* signaling pathways reporting that it is suggestive of a positive association with breast cancer.

Strengths and Limitations

Strengths of this study include the substantial sample size based on 4,700 Hispanic women (cases=2,100; controls=2,600) who completed demographic and lifestyle questionnaires and had DNA available for analysis. Data harmonization, based on variables derived from study-specific questions, was carried out and new variables

were created that used the same or the closest information possible from each original study variable to ensure consistency. The distributions of the new variables were compared and found to be very similar across these three studies providing validity to the harmonization process [8]. This process helped to rule out differential misclassification bias between ethnic groups or case-control status at different study sites.

Selection bias could be possible across study populations and countries due to differences in recruiting processes. The present study used several methods of identifying and recruiting eligible subjects including (but not limited to): mailing lists, driver's license lists, SEER registries, hospital-catchment areas, county or city-specific residence, and random-digit dialing. Some of these methods may have resulted in lower response rates, especially among H women in the US study populations [139-140]. The present study does not have the ability to measure characteristics of non-respondents and it is possible they are different than the study participants.

Although ~30% of the subjects did not provide a blood/saliva sample, the blood draw rates (or mouthwash samples) were comparable between cases (72%) and controls (76%), but were slightly higher among Hispanic (73%) and NHW participants (66%) overall, which could result in a selection bias. These issues could make this sample less representative of the general population.

Participants are being asked to remember lifestyle choices prior to diagnosis or selection which could result in recall bias. This was reduced by setting the referent year to be the same for cases and controls. Although, the present genetic association study did not show evidence of confounding by lifestyle factors.

There is also potential for genotyping errors in the lab such as cross-contamination or low concentration. This could result in drop outs, or samples that are not read during the PCR process. In order to reduce this error the concentrations of DNA samples were standardized prior to PCR and dropouts were genotyped again.

This is the first study to evaluate these particular genes in a genetically admixed population and their association with breast cancer, with the ability to compare Hispanic to NHW women, and did not have to rely on self-reported ethnicity. The STRUCTURE program determined genetic admixture, although the two were highly correlated. This program is unique in that it does not assume previous knowledge about the population; it lets the genotype define the population structure [145]. False positives may result in association studies when admixed populations with differences in incidence rates, genotype distribution, and LD between SNPs on the same locus are not adequately identified [220]. In relation to the present study, a genetic variant that is more common in NHW (high-incidence population) can appear to be associated with breast cancer in the admixed population, although no true causal association exists. By using genetic admixture as a confounder, the potential bias was controlled for risk estimates. The large sample size allowed sufficient power to test for genetic admixture as an effect modifier of the SNP and breast cancer.

Adjustment was made for multiple comparisons, reducing the potential of false positive findings, the step-down Bonferroni-Holm method is conservative, however, this does not completely remove the possibility of spurious associations [151]. The present study utilized spectral decomposition to determine the number of effective SNPs by gene before the multiple comparison adjustment. This ensured that the joint Type I error (α)

remained at 0.05 for each SNP which were jointly adjusted for multiple comparisons by gene. Another issue when adjusting for multiple comparisons is an increase in the Type II error rate, or number of false negatives that can be anticipated, which may obscure true associations [146]. Nonetheless, the present study recognizes the importance of reducing false positive results by adjusting for multiple comparisons. Very few of the above mentioned studies adjusted for multiple comparisons, mostly because they only reviewed one SNP at a time. In addition, other major findings reported were significant prior to adjusting for multiple comparisons, and should be interpreted with caution.

Because of sparse data in the literature, comparisons to previous literature were limited to only a few SNPs for *TGF-β1* (rs1982073, rs1800469), *TFG-βRI* (*6A) and *ERα* (*Pvull* (rs2234693), *XbaI* (rs9340799), rs1800132, rs3798577, rs2046210, rs851984, rs6913578) and no comparisons were possible for *RUNX* SNPs. The interpretation of findings from the present study is general and guided mostly by *in-vivo* and transgenic (genetically modified) mouse studies. While some of the SNPs have been found to modify the protein production, a true causal variant is yet to be elucidated by genetic markers located on these genes. To determine the causal variant, sequencing would have to be conducted, which was beyond the means of the present study. Using a tag-SNP approach, the present study was able to cover a large part of each gene, allowing detection of unmeasured common genetic variants, as these tagSNPs are in LD with common variants not reported in this study. Stram and colleagues explain that for the purpose of determining whether genetic variation is related to risk, it is not necessary to genotype the actual causal variant. By genotyping SNPs that also fall on the original

ancestral chromosome (gene), near the causal variant, some part of the signal of the actual causal variant will be picked up [221].

Future Direction

Replication of these findings is warranted among similarly diverse populations with larger sample sizes having available data on relevant confounders. More convincing evidence could be derived from cohort studies with incident cases. However, this may be difficult and prove to be costly. In most cases, data and resources would need to be pooled as they were in this study.

To better understand the function of these genes, the measurement of gene expression of these tagSNPs within breast tumors could provide evidence of the influence genetic variation may have on the activity, structure, and communication within and across signaling pathways [222]. An understanding of how these genes switch roles from tumor suppressors to promoters in breast cancer, which can be thought of as the concept of reprogramming gene activity in relation to tumor growth, is important for future research.

Unique findings from the present study may provide implication for further breast cancer phenotype classification. Specifically, genetic variation that attributes to ER+ and ER-tumors may be useful for examining the variance of breast cancer attributable by these specific genes and determining how they affect the tumor biology differently in Hispanic and NHW women. Understanding the biological basis of breast cancer may assist in illuminating the health disparity pertaining to differences in tumor biology and the microenvironment between Hispanic and NHW women [223]. Using these genes for further phenotype classification may prove to be a good risk predicting tool that can

specifically target differences in diverse ethnic populations. In addition, evaluating cumulative or additive effects of low-penetrance SNPs in complex genetic pathways may provide a feasible approach for prediction of risk based on common genetic variation. Using a genetic risk profile has implication for policy to make improvements in the effectiveness of population-based programs of interventions for breast cancer. In turn, policies would need to target interventions for ethnically diverse populations of various socio-economic backgrounds to further counteract the unequal burden of breast cancer [224].

Conclusion

In conclusion, the results of the present analyses provide evidence that genetic variation in *TGF-β* and *RUNX* genes may influence breast cancer risk. These associations may differ by tumor phenotype, menopausal status, and genetic admixture. These results suggest that genetic variation in these genes may explain the greater likelihood in Hispanic women for premenopausal, ER- breast cancer. This is one of the first reports that may explain these associations in Hispanic women. These associations were predominantly observed in *RUNX* genes; the present study is the first observational study to report the significant relationship between genetic variation among these genes and breast cancer risk. There is a clear implication for further tumor phenotype classification which may be useful to discriminate high- and low- risk genotypes and provide biological targets to reduce health disparities between Hispanic and NHW women.

Biologic significance of the genes is strongly suggested, although specific alleles evaluated in the present study may or may not be functional (i.e. may serve as a proxy for

other alleles). A better understanding of how *TGF-β* and *RUNX* genes can switch their role from tumor suppressor to initiator and how cross-regulatory effects of signaling pathways may contribute to breast cancer is needed. Studies evaluating a larger representation of SNPs in this complex signaling pathway will aid in validating these findings.

REFERENCES

1. *Cancer Facts and Figures 2012*. 2012, American Cancer Society: Atlanta.

2. Howlader, N., et al., *SEER Cancer Statistics Review, 1975-2009*. 2012, National Cancer Institute: Bethesda, MD.

3. Jemal, A., et al., *Global cancer statistics.* CA Cancer J Clin, 2011. **61**(2): p. 69-90.

4. Ferlay, J., et al. *Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 10 [Available from: http://globocan.iarc.fr]*. GLOBOCAN 2008 v2.0, 2008 [cited 2012 09/08].

5. Perou, C.M., et al., *Molecular portraits of human breast tumours.* Nature, 2000. **406**(6797): p. 747-52.

6. *Breast Cancer Epidemiology*, ed. C. Li. 2010: Springer.

7. Henderson, B.E., B. Ponder, and R. Ross, *Hormones, Genes, and Cancer*. 2003, New York: Oxford University Press. 450.

8. Slattery, M., et al., *Genetic variation in genes involved in hormones, inflammation, and energetic factors and breast cancer risk in an admixed population.* Carcinogenesis, 2012.

9. Blyth, K., E.R. Cameron, and J.C. Neil, *The RUNX genes: gain or loss of function in cancer.* Nat Rev Cancer, 2005. **5**(5): p. 376-87.

10. Zu, X., et al., *Transforming growth factor-beta signaling in tumor initiation, progression and therapy in breast cancer: an update.* Cell Tissue Res, 2012. **347**(1): p. 73-84.

11. Jemal, A., et al., *Annual report to the nation on the status of cancer, 1975-2001, with a special feature regarding survival.* Cancer, 2004. **101**(1): p. 3-27.

12. Wallace, T.A., D.N. Martin, and S. Ambs, *Interactions among genes, tumor biology and the environment in cancer health disparities: examining the evidence on a national and global scale.* Carcinogenesis, 2011. **32**(8): p. 1107-1121.

13. NCI. *What are cancer health disparities?* 2010 September 14, 2011]; Available from: http://crchd.cancer.gov/disparities/defined.html.

14. Fejerman, L., et al., *Genetic ancestry and risk of breast cancer among U.S. Latinas.* Cancer Res, 2008. **68**(23): p. 9723-8.

15. Fejerman, L., et al., *European ancestry is positively associated with breast cancer risk in Mexican women.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(4): p. 1074-82.

16. Knaul, F.M., et al., *Breast cancer in Mexico: a pressing priority.* Reprod Health Matters, 2008. **16**(32): p. 113-23.

17. Institute of Medicine (IOM), *Breast Cancer and The Environment: A Life Course Approach*. 2011, The National Academies Press: Washington, D.C.

18. Pike, M.C., et al., *'Hormonal' risk factors, 'breast tissue age' and the age-incidence of breast cancer.* Nature, 1983. **303**(5920): p. 767-70.

19. Rosner, B., G.A. Colditz, and W.C. Willett, *Reproductive risk factors in a prospective study of breast cancer: the Nurses' Health Study.* Am J Epidemiol, 1994. **139**(8): p. 819-35.

20. Simpson, E.R., *Sources of estrogen and their importance.* J Steroid Biochem Mol Biol, 2003. **86**(3-5): p. 225-30.

21. Harris, J., et al., *Diseases of the Breast Fourth Edition*, ed. J. Harris. 2010, Philidelphia: Wolters Kluwer.

22. Key, J., et al., *Meta-analysis of studies of alcohol and breast cancer with consideration of the methodological issues.* Cancer Causes Control, 2006. **17**(6): p. 759-70.

23. Kelsey, J.L., M.D. Gammon, and E.M. John, *Reproductive factors and breast cancer.* Epidemiol Rev, 1993. **15**(1): p. 36-47.

24. Monninkhof, E.M., et al., *Physical activity and breast cancer: a systematic review.* Epidemiology, 2007. **18**(1): p. 137-57.

25. Morimoto, L.M., et al., *Obesity, body size, and risk of postmenopausal breast cancer: the Women's Health Initiative (United States).* Cancer Causes Control, 2002. **13**(8): p. 741-51.

26. Roses, D., *Breast Cancer*. 2nd ed. 2005, Philidelphia: Elsevier.

27. Collaborative, et al., *Breast cancer and hormone replacement therapy: collaborative reanalysis of data from from 51 epidemiological studies of 52,705 women with breast cancer and 108,411 women without breast cancer.* Lancet, 1997. **15**(350): p. 1484.

28. Lacey, J.V., Jr., et al., *Breast cancer epidemiology according to recognized breast cancer risk factors in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial Cohort.* BMC Cancer, 2009. **9**: p. 84.

29. Kwan, M.L., et al., *Epidemiology of breast cancer subtypes in two prospective cohort studies of breast cancer survivors.* Breast Cancer Res, 2009. **11**(3): p. R31.

30. Li, C.I., K.E. Malone, and J.R. Daling, *Differences in breast cancer stage, treatment, and survival by race and ethnicity.* Arch Intern Med, 2003. **163**(1): p. 49-56.

31. Lee, S.A., R.K. Ross, and M.C. Pike, *An overview of menopausal oestrogen-progestin hormone therapy and breast cancer risk.* Br J Cancer, 2005. **92**(11): p. 2049-58.

32. Lee, S., et al., *Postmenopausal hormone therapy and breast cancer risk: the Multiethnic Cohort.* Int J Cancer, 2006. **118**(5): p. 1285-91.

33. *Breast cancer and hormonal contraceptives: collaborative reanalysis of individual data on 53 297 women with breast cancer and 100 239 women without breast cancer from 54 epidemiological studies. Collaborative Group on Hormonal Factors in Breast Cancer.* Lancet, 1996. **347**(9017): p. 1713-27.

34. Althuis, M.D., et al., *Breast cancers among very young premenopausal women (United States).* Cancer Causes Control, 2003. **14**(2): p. 151-60.

35. Althuis, M.D., et al., *Etiology of hormone receptor-defined breast cancer: a systematic review of the literature.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(10): p. 1558-68.

36. Wenten, M., et al., *Associations of weight, weight change, and body mass with breast cancer risk in Hispanic and non-Hispanic white women.* Ann Epidemiol, 2002. **12**(6): p. 435-4.

37. Abdel-Maksoud, M.F., et al., *Behavioral risk factors and their relationship to tumor characteristics in Hispanic and non-Hispanic white long-term breast cancer survivors.* Breast Cancer Res Treat, 2012. **131**(1): p. 169-76.

38. Fan, S., et al., *Alcohol stimulates estrogen receptor signaling in human breast cancer cell lines.* Cancer Res, 2000. **60**(20): p. 5635-9.

39. Lew, J.Q., et al., *Alcohol and risk of breast cancer by histologic type and hormone receptor status in postmenopausal women: the NIH-AARP Diet and Health Study.* Am J Epidemiol, 2009. **170**(3): p. 308-17.

40. Layde, P.M., et al., *The independent associations of parity, age at first full term pregnancy, and duration of breastfeeding with the risk of breast cancer. Cancer and Steroid Hormone Study Group.* J Clin Epidemiol, 1989. **42**(10): p. 963-73.

41. Ma, H., et al., *Reproductive factors and breast cancer risk according to joint estrogen and progesterone receptor status: a meta-analysis of epidemiological studies.* Breast Cancer Res, 2006. **8**(4): p. R43.

42. Bernstein, L., *Epidemiology of endocrine-related risk factors for breast cancer.* J Mammary Gland Biol Neoplasia, 2002. **7**(1): p. 3-15.

43. Gilliland, F.D., et al., *Reproductive risk factors for breast cancer in Hispanic and non-Hispanic white women: the New Mexico Women's Health Study.* Am J Epidemiol, 1998. **148**(7): p. 683-92.

44. Gilliland, F.D., et al., *Physical activity and breast cancer risk in hispanic and non-hispanic white women.* Am J Epidemiol, 2001. **154**(5): p. 442-50.

45. Hines, L.M., et al., *Differences in estrogen receptor subtype according to family history of breast cancer among Hispanic, but not non-Hispanic White women.* Cancer Epidemiol Biomarkers Prev, 2008. **17**(10): p. 2700-6.

46. Harper, S., et al., *Trends in area-socioeconomic and race-ethnic disparities in breast cancer incidence, stage at diagnosis, screening, mortality, and survival among women ages 50 years and over (1987-2005).* Cancer Epidemiol Biomarkers Prev, 2009. **18**(1): p. 121-31.

47. Larsen, S.B., et al., *Socioeconomic position and lifestyle in relation to breast cancer incidence among postmenopausal women: a prospective cohort study, Denmark, 1993-2006.* Cancer Epidemiol, 2011. **35**(5): p. 438-41.

48. Yost, K., et al., *Socioeconomic status and breast cancer incidence in California for different race/ethnic groups.* Cancer Causes Control, 2001. **12**(8): p. 703-11.

49. Vona-Davis, L. and D.P. Rose, *The influence of socioeconomic disparities on breast cancer tumor biology and prognosis: a review.* J Womens Health (Larchmt), 2009. **18**(6): p. 883-93.

50. Hosmer, D., Jr. and S. Lemeshow, *Applied Logistic Regression*. Probability and Mathematical Statistics, ed. J.W.a. Sons. 1989, New York: Wiley.

51. *Cancer Facts and Figures*, A.C. Society, Editor. 2011, American Cancer Society: Atlanta.

52. Hammond, M.E., et al., *American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer (unabridged version).* Arch Pathol Lab Med, 2010. **134**(7): p. e48-72.

53. Yeh, I.T. and C. Mies, *Application of immunohistochemistry to breast lesions.* Arch Pathol Lab Med, 2008. **132**(3): p. 349-58.

54. Walker, R.A., *Immunohistochemical markers as predictive tools for breast cancer.* J Clin Pathol, 2008. **61**(6): p. 689-96.

55. Wolff, A.C., et al., *American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer.* J Clin Oncol, 2007. **25**(1): p. 118-45.

56. Blows, F.M., et al., *Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a*

collaborative analysis of data for 10,159 cases from 12 studies. PLoS Med, 2010. **7**(5): p. e1000279.

57. Lander, E.S., et al., *Initial sequencing and analysis of the human genome.* Nature, 2001. **409**(6822): p. 860-921.

58. Mavaddat, N., et al., *Genetic susceptibility to breast cancer.* Mol Oncol, 2010. **4**(3): p. 174-91.

59. Antoniou, A.C., et al., *The BOADICEA model of genetic susceptibility to breast and ovarian cancers: updates and extensions.* Br J Cancer, 2008. **98**(8): p. 1457-66.

60. Birch, J.M., et al., *Relative frequency and morphology of cancers in carriers of germline TP53 mutations.* Oncogene, 2001. **20**(34): p. 4621-8.

61. Meijers-Heijboer, H., van den Ouweland, A., Klijn, J.,, et al., *Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations.* Nat Genet, 2002. **31**(55-59).

62. Renwick, A., Thompson, D., Seal, S., Kelly, P., Chagtai, T.,, et al., *ATM mutations that cause ataxiatelangiectasia are breast cancer susceptibility alleles.* Nat Genet, 2006(38): p. 873-875.

63. Seal, S., et al., *Truncating mutations in the Fanconi anemia J gene BRIP1 are low-penetrance breast cancer susceptibility alleles.* Nat Genet, 2006. **38**(11): p. 1239-41.

64. Collins, A. and I. Politopoulos, *The Genetics of Breast Cancer: Risk Factors for Disease.* The Application of Clinical Genetics, 2011(4): p. 11-19.

65. Easton, D.F., et al., *Genome-wide association study identifies novel breast cancer susceptibility loci.* Nature, 2007. **447**(7148): p. 1087-93.

66. Slattery, M.L., et al., *Replication of five GWAS-identified loci and breast cancer risk among Hispanic and non-Hispanic white women living in the Southwestern United States.* Breast Cancer Res Treat, 2011. **129**(2): p. 531-9.

67. Garcia-Closas, M., et al., *Heterogeneity of breast cancer associations with five susceptibility loci by clinical and pathological characteristics.* PLoS Genet, 2008. **4**(4): p. e1000054.

68. Long, J., et al., *Evaluation of breast cancer susceptibility loci in Chinese women.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(9): p. 2357-65.

69. Marchini, J., et al., *The effects of human population structure on large genetic association studies.* Nat Genet, 2004. **36**(5): p. 512-7.

70. Slattery, M.L. and F.A. Fitzpatrick, *Convergence of hormones, inflammation, and energy-related factors: a novel pathway of cancer etiology.* Cancer Prev Res (Phila), 2009. **2**(11): p. 922-30.

71. *ESR1 estrogen receptor 1 [Homo sapiens]* 2012, National Center for Biotechnology Information.

72. Burns, K.A. and K.S. Korach, *Estrogen receptors and human disease: an update.* Arch Toxicol, 2012. **86**(10): p. 1491-504.

73. Ali, S. and R.C. Coombes, *Estrogen receptor alpha in human breast cancer: occurrence and significance.* J Mammary Gland Biol Neoplasia, 2000. **5**(3): p. 271-81.

74. Gronemeyer, H., *Transcription activation by estrogen and progesterone receptors.* Annu Rev Genet, 1991. **25**: p. 89-123.

75. Schwabe, J.W., et al., *The crystal structure of the estrogen receptor DNA-binding domain bound to DNA: how receptors discriminate between their response elements.* Cell, 1993. **75**(3): p. 567-78.

76. Rosenfeld, M.G. and C.K. Glass, *Coregulator codes of transcriptional regulation by nuclear receptors.* J Biol Chem, 2001. **276**(40): p. 36865-8.

77. Moriarty, K., K.H. Kim, and J.R. Bender, *Minireview: estrogen receptor-mediated rapid signaling.* Endocrinology, 2006. **147**(12): p. 5557-63.

78. Deroo, B.J. and K.S. Korach, *Estrogen receptors and human disease.* J Clin Invest, 2006. **116**(3): p. 561-70.

79. Li, C.I., K.E. Malone, and J.R. Daling, *Differences in breast cancer hormone receptor status and histology by race and ethnicity among women 50 years of age and older.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(7): p. 601-7.

80. Yang, X.R., et al., *Associations of breast cancer risk factors with tumor subtypes: a pooled analysis from the Breast Cancer Association Consortium studies.* J Natl Cancer Inst, 2011. **103**(3): p. 250-63.

81. Yang, X.R., et al., *Differences in risk factors for breast cancer molecular subtypes in a population-based study.* Cancer Epidemiol Biomarkers Prev, 2007. **16**(3): p. 439-43.

82. Hill, D.A., et al., *Method of detection and breast cancer survival disparities in Hispanic women.* Cancer Epidemiol Biomarkers Prev, 2010. **19**(10): p. 2453-60.

83. Hines, L.M., et al., *Comparative analysis of breast cancer risk factors among Hispanic and non-Hispanic white women.* Cancer, 2010. **116**(13): p. 3215-23.

84. Li, C.I., K.E. Malone, and J.R. Daling, *Differences in Breast Cancer Stage, Treatment, and Survival by Race and Ethnicity.* Arch Intern Med, 2003. **163**: p. 49-56.

85. Shavers, V.L., L.C. Harlan, and J.L. Stevens, *Racial/ethnic variation in clinical presentation, treatment, and survival among breast cancer patients under age 35.* Cancer, 2003. **97**(1): p. 134-47.

86. Martinez, M.E., et al., *Breast cancer among Hispanic and non-Hispanic White women in Arizona.* J Health Care Poor Underserved, 2007. **18**(4 Suppl): p. 130-45.

87. Elliott, R.L. and G.C. Blobe, *Role of transforming growth factor Beta in human cancer.* J Clin Oncol, 2005. **23**(9): p. 2078-93.

88. Scollen, S., et al., *TGF-beta signaling pathway and breast cancer susceptibility.* Cancer Epidemiol Biomarkers Prev, 2011. **20**(6): p. 1112-9.

89. Heldin, C.H., K. Miyazono, and P. ten Dijke, *TGF-beta signalling from cell membrane to nucleus through SMAD proteins.* Nature, 1997. **390**(6659): p. 465-71.

90. *TGFB1 transforming growth factor, beta 1 [Homo sapiens]* 2012, National Center for Biotechnology Information.

91. Shi, Y. and J. Massague, *Mechanisms of TGF-beta signaling from cell membrane to the nucleus.* Cell, 2003. **113**(6): p. 685-700.

92. *TGFBR1 transforming growth factor, beta receptor 1 [Homo sapiens]* 2012, National Center for Biotechnology Information.

93. Massague, J., *TGFbeta in Cancer.* Cell, 2008. **134**(2): p. 215-30.

94. Pierce, D.F., Jr., et al., *Mammary tumor suppression by transforming growth factor beta 1 transgene expression.* Proc Natl Acad Sci U S A, 1995. **92**(10): p. 4254-8.

95. Moustakas, A., et al., *Mechanisms of TGF-beta signaling in regulation of cell growth and differentiation.* Immunol Lett, 2002. **82**(1-2): p. 85-91.

96. Levy, L. and C.S. Hill, *Alterations in components of the TGF-beta superfamily signaling pathways in human cancer.* Cytokine Growth Factor Rev, 2006. **17**(1-2): p. 41-58.

97. Serra, R. and M.R. Crowley, *Mouse models of transforming growth factor beta impact in breast development and cancer.* Endocr Relat Cancer, 2005. **12**(4): p. 749-60.

98. Ewen, M.E., et al., *p53-dependent repression of CDK4 translation in TGF-beta-induced G1 cell-cycle arrest.* Genes Dev, 1995. **9**(2): p. 204-17.

99. Li, Q.L., et al., *Causal relationship between the loss of RUNX3 expression and gastric cancer.* Cell, 2002. **109**(1): p. 113-24.

100. Pertovaara, L., et al., *Vascular endothelial growth factor is induced in response to transforming growth factor-beta in fibroblastic and epithelial cells.* J Biol Chem, 1994. **269**(9): p. 6271-4.

101. Cox, D.G., et al., *TGFB1 and TGFBR1 polymorphisms and breast cancer risk in the Nurses' Health Study.* BMC Cancer, 2007. **7**: p. 175.

102. Kaklamani, V.G., et al., *Combined genetic assessment of transforming growth factor-beta signaling pathway variants may predict breast cancer risk.* Cancer Res, 2005. **65**(8): p. 3454-61.

103. Qi, X., et al., *Transforming growth factor-beta1 polymorphisms and breast cancer risk: a meta-analysis based on 27 case-control studies.* Breast Cancer Res Treat, 2010. **122**(1): p. 273-9.

104. Cameron, E.R. and J.C. Neil, *The Runx genes: lineage-specific oncogenes and tumor suppressors.* Oncogene, 2004. **23**(24): p. 4308-14.

105. Ito, Y., *Oncogenic potential of the RUNX gene family: 'overview'.* Oncogene, 2004. **23**(24): p. 4198-208.

106. Lund, A.H. and M. van Lohuizen, *RUNX: a trilogy of cancer genes.* Cancer Cell, 2002. **1**(3): p. 213-5.

107. Janes, K.A., *RUNX1 and its understudied role in breast cancer.* Cell Cycle, 2011. **10**(20): p. 3461-5.

108. Miyazono, K., S. Maeda, and T. Imamura, *Coordinate regulation of cell growth and differentiation by TGF-beta superfamily and Runx proteins.* Oncogene, 2004. **23**(24): p. 4232-7.

109. Ito, Y. and K. Miyazono, *RUNX transcription factors as key targets of TGF-beta superfamily signaling.* Curr Opin Genet Dev, 2003. **13**(1): p. 43-7.

110. *RUNX1 runt-related transcription factor 1 [Homo sapiens]* 2012, National Center for Biomedical Information.

111. Hiebert, S.W., B. Lutterbach, and J. Amann, *Role of co-repressors in transcriptional repression mediated by the t(8;21), t(16;21), t(12;21), and inv(16) fusion proteins.* Curr Opin Hematol, 2001. **8**(4): p. 197-200.

112. Look, A.T., *Oncogenic transcription factors in the human acute leukemias.* Science, 1997. **278**(5340): p. 1059-64.

113. Silva, F.P., et al., *Identification of RUNX1/AML1 as a classical tumor suppressor gene.* Oncogene, 2003. **22**(4): p. 538-47.

114. Uhlen, M., et al., *A human protein atlas for normal and cancer tissues based on antibody proteomics.* Mol Cell Proteomics, 2005. **4**(12): p. 1920-32.

115. Wang, L., J.S. Brugge, and K.A. Janes, *Intersection of FOXO- and RUNX1-mediated gene expression programs in single breast epithelial cells during morphogenesis and tumor progression.* Proc Natl Acad Sci U S A, 2011. **108**(40): p. E803-12.

116. *RUNX2 runt-related transcription factor 2 [Homo sapiens]* 2012, National Center for Biomedical Information.

117. Guise, T.A., *Parathyroid hormone-related protein and bone metastases.* Cancer, 1997. **80**(8 Suppl): p. 1572-80.

118. Guise, T.A., et al., *Evidence for a causal role of parathyroid hormone-related protein in the pathogenesis of human breast cancer-mediated osteolysis.* J Clin Invest, 1996. **98**(7): p. 1544-9.

119. Pratap, J., et al., *Runx2 transcriptional activation of Indian Hedgehog and a downstream bone metastatic pathway in breast cancer cells.* Cancer Res, 2008. **68**(19): p. 7795-802.

120. Zelzer, E., et al., *Tissue specific regulation of VEGF expression during bone development requires Cbfa1/Runx2.* Mech Dev, 2001. **106**(1-2): p. 97-106.

121. Onodera, Y., et al., *Runx2 in human breast carcinoma: its potential roles in cancer progression.* Cancer Sci, 2010. **101**(12): p. 2670-5.

122. *RUNX3 runt-related transcription factor 3 [Homo sapiens]* 2012, National Center for Biomedical Information.

123. Subramaniam, M.M., et al., *Molecular pathology of RUNX3 in human carcinogenesis.* Biochim Biophys Acta, 2009. **1796**(2): p. 315-31.

124. Chuang, L.S. and Y. Ito, *RUNX3 is multifunctional in carcinogenesis of multiple solid tumors.* Oncogene, 2010. **29**(18): p. 2605-15.

125. Lau, Q.C., et al., *RUNX3 is frequently inactivated by dual mechanisms of protein mislocalization and promoter hypermethylation in breast cancer.* Cancer Res, 2006. **66**(13): p. 6512-20.

126. Fujii, S., et al., *Enhancer of zeste homologue 2 (EZH2) down-regulates RUNX3 by increasing histone H3 methylation.* J Biol Chem, 2008. **283**(25): p. 17324-32.

127. Wolff, E.M., et al., *RUNX3 methylation reveals that bladder tumors are older in patients with a history of smoking.* Cancer Res, 2008. **68**(15): p. 6208-14.

128. Chen, C.R., et al., *E2F4/5 and p107 as Smad cofactors linking the TGFbeta receptor to c-myc repression.* Cell, 2002. **110**(1): p. 19-32.

129. Matsuda, T., et al., *Cross-talk between transforming growth factor-beta and estrogen receptor signaling through Smad3.* J Biol Chem, 2001. **276**(46): p. 42908-14.

130. Ito, I., et al., *Estrogen inhibits transforming growth factor beta signaling by promoting Smad2/3 degradation.* J Biol Chem, 2010. **285**(19): p. 14747-55.

131. Ren, Y., et al., *Dual effects of TGF-beta on ERalpha-mediated estrogenic transcriptional activity in breast cancer.* Mol Cancer, 2009. **8**: p. 111.

132. Wu, L., et al., *Smad4 as a transcription corepressor for estrogen receptor alpha.* J Biol Chem, 2003. **278**(17): p. 15192-200.

133. Band, A.M. and M. Laiho, *Crosstalk of TGF-beta and estrogen receptor signaling in breast cancer.* J Mammary Gland Biol Neoplasia, 2011. **16**(2): p. 109-15.

134. Araki, S., et al., *TGF-beta1-induced expression of human Mdm2 correlates with late-stage metastatic breast cancer.* J Clin Invest, 2010. **120**(1): p. 290-302.

135. Bierie, B., et al., *Abrogation of TGF-beta signaling enhances chemokine production and correlates with prognosis in human breast cancer.* J Clin Invest, 2009. **119**(6): p. 1571-82.

136. Huang, B., et al., *RUNX3 acts as a tumor suppressor in breast cancer by targeting estrogen receptor alpha.* Oncogene, 2011.

137. Key, T.J., P.K. Verkasalo, and E. Banks, *Epidemiology of breast cancer.* Lancet Oncol, 2001. **2**(3): p. 133-40.

138. Lichtenstein, P., et al., *Environmental and heritable factors in the causation of cancer--analyses of cohorts of twins from Sweden, Denmark, and Finland.* N Engl J Med, 2000. **343**(2): p. 78-85.

139. Slattery, M.L., et al., *Body size, weight change, fat distribution and breast cancer risk in Hispanic and non-Hispanic white women.* Breast Cancer Res Treat, 2007. **102**(1): p. 85-101.

140. John, E.M., P.L. Horn-Ross, and J. Koo, *Lifetime physical activity and breast cancer risk in a multiethnic population: the San Francisco Bay area breast cancer study.* Cancer Epidemiol Biomarkers Prev, 2003. **12**(11 Pt 1): p. 1143-52.

141. Angeles-Llerenas, A., et al., *Moderate physical activity and breast cancer risk: the effect of menopausal status.* Cancer Causes Control, 2010. **21**(4): p. 577-86.

142. Buechley, R.W., *In Generally Useful Ethnic Search System: GUESS*, C.R.a.T. Center, Editor. 1976, University of New Mexico: Albuquerque, NM.

143. Word, D.L. and R.C. Perkins, *Building a Spanish Surname list for the 1990's-A new approach to an old problem* in *Technical Working Paper No.13*. 1996, Population Division, U.S. Bureau of the Census: Washington, DC. p. pp.1-25.

144. Collins-Schramm, H.E., et al., *Mexican American ancestry-informative markers: examination of population structure and marker characteristics in European Americans, Mexican Americans, Amerindians and Asians.* Hum Genet, 2004. **114**(3): p. 263-71.

145. Pritchard, J.K., M. Stephens, and P. Donnelly, *Inference of population structure using multilocus genotype data.* Genetics, 2000. **155**(2): p. 945-59.

146. Laird, N. and C. Lange, *The Fundamentals of Modern Statistical Genetics*. Statistics for Biology and Health, ed. M. Gail, et al. 2011, New York, NY: Springer. 223 pages.

147. Benjamini, Y., and Hochberg, Y., *Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing.* Journal of the Royal Statistical Society, 1995. **B**(57): p. 289-300.

148. Bursac, Z., et al., *Purposeful selection of variables in logistic regression.* Source Code Biol Med, 2008. **3**: p. 17.

149. Dubin, N. and B.S. Pasternack, *Risk assessment for case-control subgroups by polychotomous logistic regression.* Am J Epidemiol, 1986. **123**(6): p. 1101-17.

150. Dai, J., et al., *Breast cancer risk assessment with five independent genetic variants and two risk factors in Chinese women.* Breast Cancer Res, 2012. **14**(1): p. R17.

151. Holm, S., *A simple sequentially rejective multiple test procedure.* Scand. J. Stat, 1979. **6**: p. 65-70.

152. Nyholt, D.R., *A simple correction for multiple testing for single-nucleotide polymorphisms in linkage disequilibrium with each other.* Am J Hum Genet, 2004. **74**(4): p. 765-9.

153. Li, J. and L. Ji, *Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix.* Heredity (Edinb), 2005. **95**(3): p. 221-7.

154. Purcell, S., S.S. Cherny, and P.C. Sham, *Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits.* Bioinformatics, 2003. **19**(1): p. 149-50.

155. Akhurst, R.J., *TGF beta signaling in health and disease.* Nat Genet, 2004. **36**(8): p. 790-2.

156. Cox, A., et al., *A common coding variant in CASP8 is associated with breast cancer risk.* Nat Genet, 2007. **39**(3): p. 352-8.

157. Le Marchand, L., et al., *T29C polymorphism in the transforming growth factor beta1 gene and postmenopausal breast cancer risk: the Multiethnic Cohort Study.* Cancer Epidemiol Biomarkers Prev, 2004. **13**(3): p. 412-5.

158. Hishida, A., et al., *Transforming growth factor B1 T29C polymorphism and breast cancer risk in Japanese women.* Breast Cancer, 2003. **10**(1): p. 63-9.

159. Dunning, A.M., et al., *A transforming growth factorbeta1 signal peptide variant increases secretion in vitro and is associated with increased incidence of invasive breast cancer.* Cancer Res, 2003. **63**(10): p. 2610-5.

160. Shin, A., et al., *Genetic polymorphisms of the transforming growth factor-beta1 gene and breast cancer risk: a possible dual role at different cancer stages.* Cancer Epidemiol Biomarkers Prev, 2005. **14**(6): p. 1567-70.

161. The MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk, *Polymorphisms in the BRCA1 and ABCB1 genes modulate menopausal hormone therapy associated breast cancer risk in postmenopausal women.* Breast Cancer Res Treat, 2010. **120**(3): p. 727-36.

162. Jin, Q., et al., *Polymorphisms and haplotype structures in genes for transforming growth factor beta1 and its receptors in familial and unselected breast cancers.* Int J Cancer, 2004. **112**(1): p. 94-9.

163. Colleran, G., et al., *The TGFBR1*6A/9A polymorphism is not associated with differential risk of breast cancer.* Breast Cancer Res Treat, 2010. **119**(2): p. 437-42.

164. Zhang, H.T., et al., *Is TGFBR1*6A really associated with increased risk of cancer?* J Clin Oncol, 2005. **23**(30): p. 7743-4; author reply 7744-6.

165. Baxter, S.W., et al., *Transforming growth factor beta receptor 1 polyalanine polymorphism and exon 5 mutation analysis in breast and ovarian cancer.* Cancer Epidemiol Biomarkers Prev, 2002. **11**(2): p. 211-4.

166. Chen, T., et al., *Int7G24A variant of transforming growth factor-beta receptor type I is associated with invasive breast cancer.* Clin Cancer Res, 2006. **12**(2): p. 392-7.

167. Feigelson, H.S., et al., *Transforming growth factor beta receptor type I and transforming growth factor beta1 polymorphisms are not associated with postmenopausal breast cancer.* Cancer Epidemiol Biomarkers Prev, 2006. **15**(6): p. 1236-7.

168. Chen, T., et al., *An intronic variant of the TGFBR1 gene is associated with carcinomas of the kidney and bladder.* Int J Cancer, 2004. **112**(3): p. 420-5.

169. Pasche, B., et al., *Constitutively decreased TGFBR1 allelic expression is a common finding in colorectal cancer and is associated with three TGFBR1 SNPs.* J Exp Clin Cancer Res, 2010. **29**: p. 57.

170. Wang, Y.Q., et al., *Association between TGFBR1 polymorphisms and cancer risk: a meta-analysis of 35 case-control studies.* PLoS One, 2012. **7**(8): p. e42899.

171. Schirmer, M.A., et al., *Bioinformatic and functional analysis of TGFBR1 polymorphisms.* Pharmacogenet Genomics, 2009. **19**(4): p. 249-59.

172. Grainger, D.J., et al., *Genetic control of the circulating concentration of transforming growth factor type beta1.* Hum Mol Genet, 1999. **8**(1): p. 93-7.

173. Yokota, M., et al., *Association of a T29-->C polymorphism of the transforming growth factor-beta1 gene with genetic susceptibility to myocardial infarction in Japanese.* Circulation, 2000. **101**(24): p. 2783-7.

174. Wei, B.B., et al., *TGFbeta1 T29C polymorphism and cancer risk: a meta-analysis based on 40 case-control studies.* Cancer Genet Cytogenet, 2010. **196**(1): p. 68-75.

175. Woo, S.U., et al., *Association of a TGF-beta1 gene -509 C/T polymorphism with breast cancer risk: a meta-analysis.* Breast Cancer Res Treat, 2010. **124**(2): p. 481-5.

176. Niu, W., et al., *Association of TGFB1 -509 C>T polymorphism with breast cancer: evidence from a meta-analysis involving 23,579 subjects.* Breast Cancer Res Treat, 2010. **124**(1): p. 243-9.

177. Jakubowska, A., et al., *Do BRCA1 modifiers also affect the risk of breast cancer in non-carriers?* Eur J Cancer, 2009. **45**(5): p. 837-42.

178. Slattery, M.L., et al., *Associations between genetic variation in RUNX1, RUNX2, RUNX3, MAPK1 and eIF4E and riskof colon and rectal cancer: additional support for a TGF-beta-signaling pathway.* Carcinogenesis, 2011. **32**(3): p. 318-26.

179. Becherini, L., et al., *Evidence of a linkage disequilibrium between polymorphisms in the human estrogen receptor alpha gene and their relationship to bone mass variation in postmenopausal Italian women.* Hum Mol Genet, 2000. **9**(13): p. 2043-50.

180. Dunning, A.M., et al., *Association of ESR1 gene tagging SNPs with breast cancer risk.* Hum Mol Genet, 2009. **18**(6): p. 1131-9.

181. Einarsdottir, K., et al., *ESR1 and EGF genetic variation in relation to breast cancer risk and survival.* Breast Cancer Res, 2008. **10**(1): p. R15.

182. Fernandez, L.P., et al., *Estrogen and progesterone receptor gene polymorphisms and sporadic breast cancer risk: a Spanish case-control study.* Int J Cancer, 2006. **119**(2): p. 467-71.

183. Gold, B., et al., *Estrogen receptor genotypes and haplotypes associated with breast cancer risk.* Cancer Res, 2004. **64**(24): p. 8891-900.

184. Gonzalez-Zuloeta Ladd, A.M., et al., *Estrogen receptor alpha polymorphisms and postmenopausal breast cancer risk.* Breast Cancer Res Treat, 2008. **107**(3): p. 415-9.

185. Wang, J., et al., *Estrogen receptor alpha haplotypes and breast cancer risk in older Caucasian women.* Breast Cancer Res Treat, 2007. **106**(2): p. 273-80.

186. Onland-Moret, N.C., et al., *The estrogen receptor alpha gene and breast cancer risk (The Netherlands).* Cancer Causes Control, 2005. **16**(10): p. 1195-202.

187. Slattery, M.L., et al., *ESR1, AR, body size, and breast cancer risk in Hispanic and non-Hispanic white women living in the Southwestern United States.* Breast Cancer Res Treat, 2007. **105**(3): p. 327-35.

188. Li, N., et al., *Potentially functional polymorphisms in ESR1 and breast cancer risk: a meta-analysis.* Breast Cancer Res Treat, 2010. **121**(1): p. 177-84.

189. Anghel, A., et al., *Estrogen receptor alpha polymorphisms: correlation with clinicopathological parameters in breast cancer.* Neoplasma, 2010. **57**(4): p. 306-15.

190. Gallicchio, L., et al., *Polymorphisms in estrogen-metabolizing and estrogen receptor genes and the risk of developing breast cancer among a cohort of women with benign breast disease.* BMC Cancer, 2006. **6**: p. 173.

191. Han, J., et al., *Genetic variants of 6q25 and breast cancer susceptibility: a two-stage fine mapping study in a Chinese population.* Breast Cancer Res Treat, 2011. **129**(3): p. 901-7.

192. Siddig, A., et al., *Estrogen receptor alpha gene polymorphism and breast cancer.* Ann N Y Acad Sci, 2008. **1138**: p. 95-107.

193. Wedren, S., et al., *Oestrogen receptor alpha gene haplotype and postmenopausal breast cancer risk: a case control study.* Breast Cancer Res, 2004. **6**(4): p. R437-49.

194. Tapper, W., et al., *The influence of genetic variation in 30 selected genes on the clinical characteristics of early onset breast cancer.* Breast Cancer Res, 2008. **10**(6): p. R108.

195. Zheng, W., et al., *Genome-wide association study identifies a new breast cancer susceptibility locus at 6q25.1.* Nat Genet, 2009. **41**(3): p. 324-8.

196. Cai, Q., et al., *Replication and functional genomic analyses of the breast cancer susceptibility locus at 6q25.1 generalize its importance in women of chinese, Japanese, and European ancestry.* Cancer Res, 2011. **71**(4): p. 1344-55.

197. The MARIE-GENICA Consortium on Genetic Susceptibility for Menopausal Hormone Therapy Related Breast Cancer Risk, *Polymorphisms in genes of the steroid receptor superfamily modify postmenopausal breast cancer risk*

associated with menopausal hormone therapy. Int J Cancer, 2010. **126: 2935–2946. doi: 10.1002/ijc.24892**.

198. Oesterreich, S., et al., *Estrogen-mediated down-regulation of E-cadherin in breast cancer cells.* Cancer Res, 2003. **63**(17): p. 5203-8.

199. Burcham, P.C., *Internal hazards: baseline DNA damage by endogenous products of normal metabolism.* Mutat Res, 1999. **443**(1-2): p. 11-36.

200. Yager, J.D. and J.G. Liehr, *Molecular mechanisms of estrogen carcinogenesis.* Annu Rev Pharmacol Toxicol, 1996. **36**: p. 203-32.

201. Lavigne, J.A., et al., *An association between the allele coding for a low activity variant of catechol-O-methyltransferase and the risk for breast cancer.* Cancer Res, 1997. **57**(24): p. 5493-7.

202. Kato, S., et al., *Activation of the estrogen receptor through phosphorylation by mitogen-activated protein kinase.* Science, 1995. **270**(5241): p. 1491-4.

203. Iwase, H., et al., *Genetic and epigenetic alterations of the estrogen receptor gene and hormone independence in human breast cancer.* Oncology, 1998. **55 Suppl 1**: p. 11-6.

204. Derynck, R., R.J. Akhurst, and A. Balmain, *TGF-beta signaling in tumor suppression and cancer progression.* Nat Genet, 2001. **29**(2): p. 117-29.

205. Yue, J. and K.M. Mulder, *Requirement of Ras/MAPK pathway activation by transforming growth factor beta for transforming growth factor beta 1 production in a smad-dependent pathway.* J Biol Chem, 2000. **275**(45): p. 35656.

206. Yue, J. and K.M. Mulder, *Activation of the mitogen-activated protein kinase pathway by transforming growth factor-beta.* Methods Mol Biol, 2000. **142**: p. 125-31.

207. de Jong, J.S., et al., *Expression of growth factors, growth-inhibiting factors, and their receptors in invasive breast cancer. II: Correlations with proliferation and angiogenesis.* J Pathol, 1998. **184**(1): p. 53-7.

208. Saito, H., et al., *The expression of transforming growth factor-beta1 is significantly correlated with the expression of vascular endothelial growth factor and poor prognosis of patients with advanced gastric carcinoma.* Cancer, 1999. **86**(8): p. 1455-62.

209. Chen, W., et al., *Targets of genome copy number reduction in primary breast cancers identified by integrative genomics.* Genes Chromosomes Cancer, 2007. **46**(3): p. 288-301.

210. Saha Roy, S. and R.K. Vadlamudi, *Role of estrogen receptor signaling in breast cancer metastasis.* Int J Breast Cancer, 2012. **2012**: p. 654698.

211. Fuqua, S.A., *The role of estrogen receptors in breast cancer metastasis.* J Mammary Gland Biol Neoplasia, 2001. **6**(4): p. 407-17.

212. Slattery, M.L., et al., *IL6, aspirin, nonsteroidal anti-inflammatory drugs, and breast cancer risk in women living in the southwestern United States.* Cancer Epidemiol Biomarkers Prev, 2007. **16**(4): p. 747-55.

213. Rollison, D.E., et al., *Serum insulin-like growth factor (IGF)-1 and IGF binding protein-3 in relation to breast cancer among Hispanic and white, non-Hispanic women in the US Southwest.* Breast Cancer Res Treat, 2010. **121**(3): p. 661-9.

214. Connor, A., et al., *ADRB2 G-G haplotype associated with breast cancer risk among Hispanic and non-Hispanic white women: interaction with type 2 diabetes and obesity.* Cancer Causes Control, 2012. **23**(10): p. 1653-63.

215. Slattery, M.L., et al., *Genetic variation in bone morphogenetic proteins and breast cancer risk in hispanic and non-hispanic white women: The breast cancer health disparities study.* Int J Cancer, 2012.

216. Chimge, N.O. and B. Frenkel, *The RUNX family in breast cancer: relationships with estrogen signaling.* Oncogene, 2012.

217. Stender, J.D., et al., *Genome-wide analysis of estrogen receptor alpha DNA binding and tethering mechanisms identifies Runx1 as a novel tethering factor in receptor-mediated transcriptional activation.* Mol Cell Biol, 2010. **30**(16): p. 3943-55.

218. Chimge, N.O., et al., *Opposing effects of Runx2 and estradiol on breast cancer cell proliferation: in vitro identification of reciprocally regulated gene signature related to clinical letrozole responsiveness.* Clin Cancer Res, 2012. **18**(3): p. 901-11.

219. Posthuma, D., et al., *A note on permutation tests for genetic association analysis of quantitative traits when variances are heterogeneous.* Genet Epidemiol, 2009. **33**(8): p. 710-6.

220. Sweeney, C., et al., *Genetic admixture among Hispanics and candidate gene polymorphisms: potential for confounding in a breast cancer study?* Cancer Epidemiol Biomarkers Prev, 2007. **16**(1): p. 142-50.

221. Stram, D.O., *Tag SNP selection for association studies.* Genet Epidemiol, 2004. **27**(4): p. 365-74.

222. Williams, R.B., et al., *The influence of genetic variation on gene expression.* Genome Res, 2007. **17**(12): p. 1707-16.

223. Lee, C.I. and A. Naeim, *Health disparities from future genetic research efforts: breast cancer as a case study.* J Natl Med Assoc, 2012. **104**(7-8): p. 390-1.

224. Pharoah, P.D., et al., *Polygenic susceptibility to breast cancer and implications for prevention.* Nat Genet, 2002. **31**(1): p. 33-6.

APPENDIX A:

SUPPLEMENTAL FIGURES

Figure 2b:
2 DF Power Analysis Curve-TGF-β Signaling SNPs
n=7,733 (Cases=3,524; Controls=4,209)

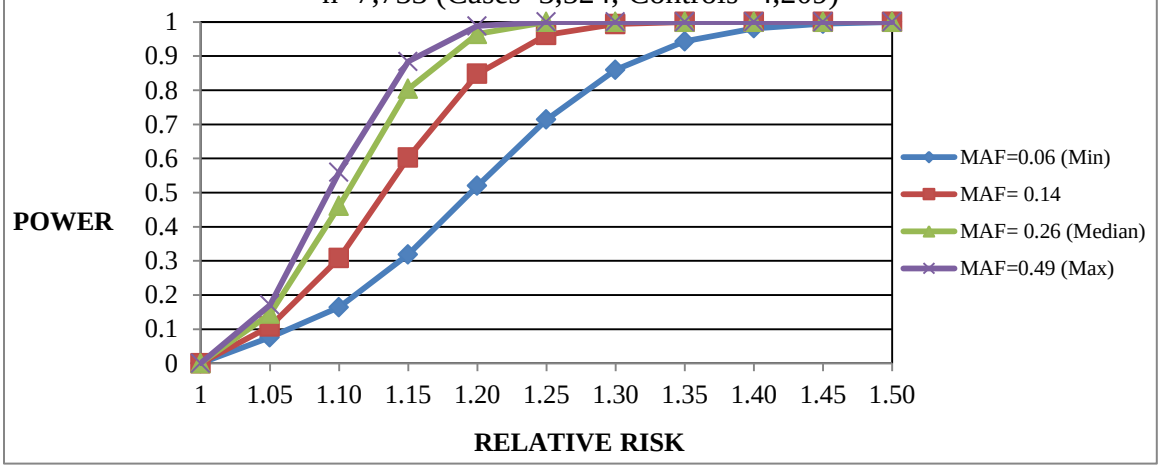


Figure 3b: Hispanic
2 DF Power Analysis Curve-TGF-β Signaling SNPs
n=4,703 (Cases=2,093; Controls=2,610)

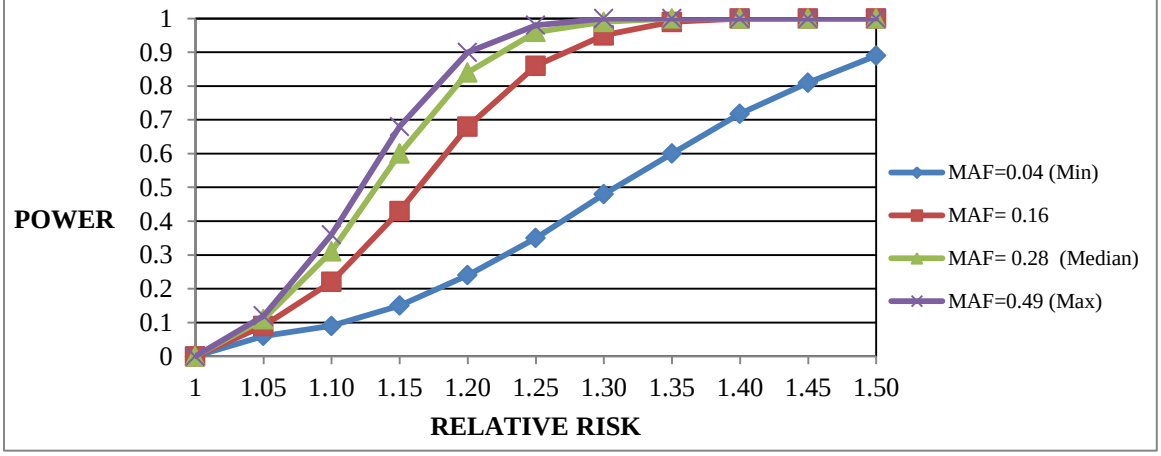


Figure 4b: non-Hispanic White
2 DF Power Analysis Curve-TGF-β Signaling SNPs
n=3,031 (Cases=1,431; Controls=1,600)

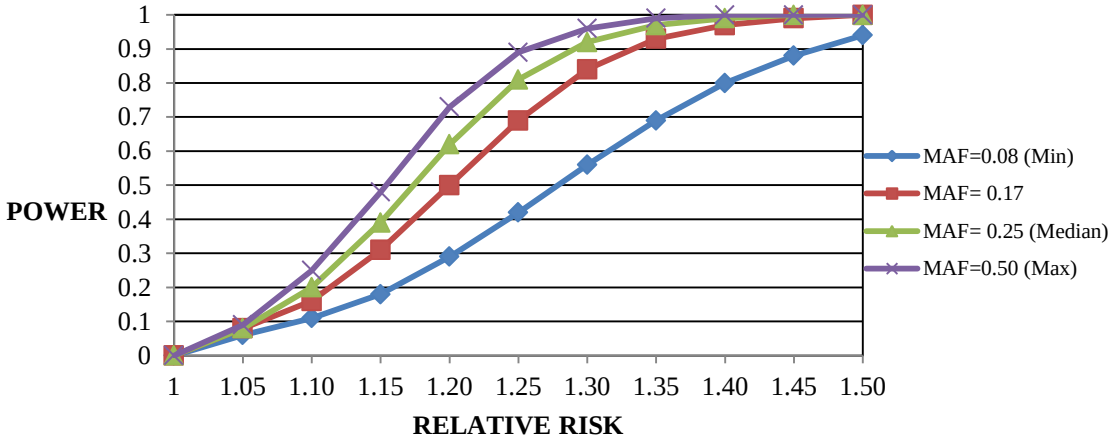


Figure 5b:
2 DF Power Analysis Curve-ERα Signaling SNPs
n=1,409 (Cases=694; Controls=715)

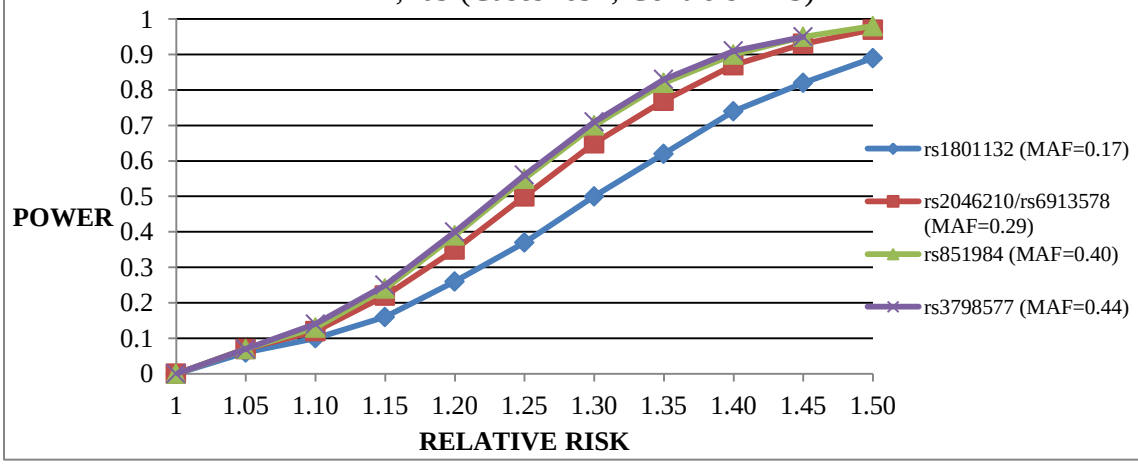


Figure 6b: Hispanic
2 DF Power Analysis Curve-ERα Signaling SNPs
n=482 (Cases=242; Controls=240)

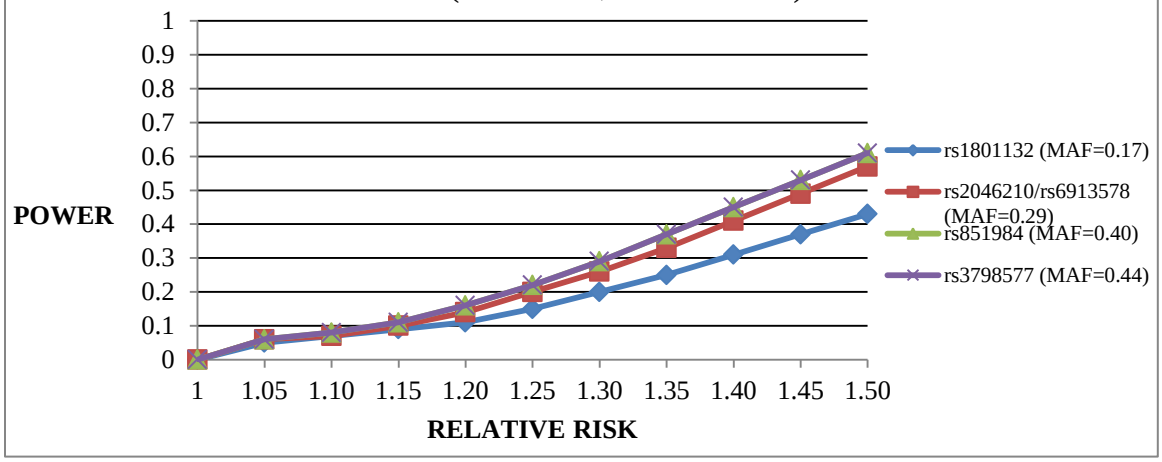
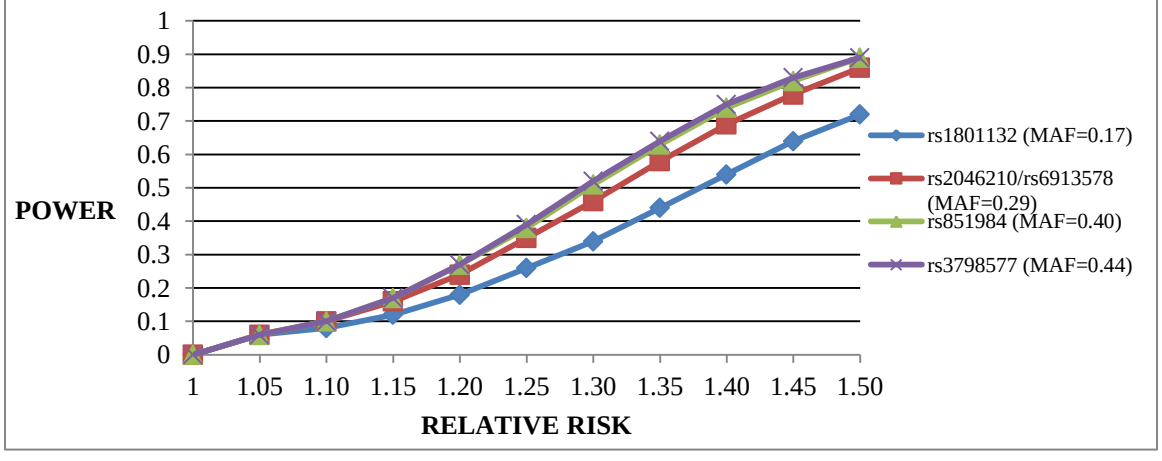


Figure 7b: non-Hispanic White
2 DF Power Analysis Curve-ERα Signaling SNPs
n=927 (Cases=452; Controls=475)



APPENDIX B:

SUPPLEMENTAL TABLES

Table 7. TGF-β signaling pathway: genotype distributions, BCHD population, stratified by self-reported race and case-control status

		Total (n=7,733)			Non-Hispanic White (n=3,030)			Hispanic (n=4,703)			
		Cases (n=3524)	Controls (n=4209)	p[a]	Cases (n=1431)	Controls (n=1599)	p[a]	Cases (n=2093)	Controls (n=2610)	p[a]	p[b]
Gene/SNP	**Genotype[c]**	***N (%)***	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
RUNX1 *(n=8)*											
rs11701453	GG	2395 (68.0)	2917 (69.3)		913 (63.9)	1033 (64.6)		1482 (70.8)	1884 (72.2)		
	GC	1030 (29.2)	1169 (27.8)	0.37	459 (32.1)	498 (31.1)	0.82	571 (27.3)	671 (25.7)	0.46	<.0001
	CC	97 (2.8)	124 (2.9)		58 (4.0)	69 (4.3)		39 (1.9)	55 (2.1)		
rs1474479	GG	2014 (57.2)	2432 (57.8)		586 (40.9)	638 (39.9)		1428 (68.2)	1794 (68.7)		
	GA	1237 (35.1)	1443 (34.3)	0.80	646 (45.2)	717 (44.8)	0.31	591 (28.4)	726 (27.8)	0.71	<.0001
	AA	272 (7.7)	335 (7.9)		198 (13.9)	245 (15.3)		74 (3.5)	90 (3.5)		
rs1883066	GG	2856 (81.0)	3471 (82.5)		1081 (75.5)	1229 (76.9)		1775 (84.8)	2242 (85.9)		
	GC	633 (17.9)	696 (16.5)	0.14	330 (23.1)	341 (21.3)	0.59	303 (14.5)	355 (13.6)	0.23	<.0001
	CC	35 (1.0)	42 (1.0)		20 (1.4)	29 (1.8)		15 (0.7)	13 (0.5)		
rs2252585	TT	1456 (41.3)	1697 (40.3)		763 (53.3)	841 (52.6)		693 (33.1)	856 (32.8)		
	TC	1584 (45.0)	1889 (44.9)	0.18	576 (40.3)	645 (40.3)	0.52	1008 (48.2)	1244 (47.7)	0.57	<.0001
	CC	483 (13.7)	624 (14.8)		92 (6.4)	114 (7.1)		391 (18.7)	510 (19.5)		
rs2268288	TT	2500 (70.9)	3027 (71.9)		890 (62.2)	1014 (63.4)		1610 (77.0)	2013 (77.1)		
	TC	915 (26.0)	1081 (25.7)	0.18	469 (32.8)	520 (32.5)	0.32	446 (21.3)	561 (21.5)	0.71	<.0001
	CC	108 (3.1)	102 (2.4)		72 (5.0)	66 (4.1)		36 (1.7)	36 (1.4)		
rs7279123	CC	2236 (63.7)	2629 (62.7)	0.31	818 (57.5)	870 (54.7)	0.15	1418 (67.9)	1759 (67.6)	0.63	<.0001
	CT	1109 (31.6)	1361 (32.4)		506 (35.6)	606 (38.1)		603 (28.9)	755 (28.9)		
	TT	163 (4.7)	206 (4.9)		98 (6.9)	116 (7.3)		65 (3.1)	90 (3.5)		
rs7279383	CC	2647 (75.1)	3099 (73.6)		977 (68.3)	1051 (65.7)		1670 (79.8)	2048 (78.5)		
	CG	812 (23.1)	1033 (24.5)	0.17	424 (29.7)	501 (31.3)	0.07	388 (18.5)	532 (20.4)	0.54	<.0001
	GG	64 (1.8)	77 (1.8)		29 (2.0)	47 (2.9)		35 (1.7)	30 (1.2)		
rs8127225	TT	2117 (60.1)	2591 (61.6)	0.26	1054 (73.7)	1205 (75.4)	0.44	1063 (50.8)	1386 (53.2)	0.13	<.0001

		Total (n=7,733)			Non-Hispanic White (n=3,030)			Hispanic (n=4,703)			
		Cases (n=3524)	Controls (n=4209)	p[a]	Cases (n=1431)	Controls (n=1599)	p[a]	Cases (n=2093)	Controls (n=2610)	p[a]	p[b]
Gene/SNP	**Genotype[c]**	***N (%)***	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
	TC	1215 (34.5)	1383 (32.9)		355 (24.8)	364 (22.8)		860 (41.1)	1019 (39.1)		
	CC	192 (5.4)	230 (5.5)		22 (1.5)	30 (1.9)		170 (8.1)	200 (7.7)		
RUNX2 ***(n=17)***											
rs10948238	CC	1245 (35.4)	1516 (36.0)		487 (34.1)	574 (35.9)		758 (36.2)	942 (36.1)		
	CT	1705 (48.4)	2086 (49.6)	0.11	692 (48.4)	800 (50.0)	0.04	1013 (48.4)	1286 (49.3)	0.74	0.23
	TT	571 (16.2)	605 (14.4)		250 (17.5)	225 (14.1)		321 (15.3)	380 (14.6)		
rs1200428	CC	1932 (54.8)	2316 (55.0)		825 (57.7)	953 (59.6)		1107 (52.9)	1363 (52.2)		
	CA	1370 (38.9)	1629 (38.7)	0.89	533 (37.2)	585 (36.6)	0.14	837 (40.0)	1044 (40.0)	0.47	<.0001
	AA	222 (6.3)	265 (6.3)		73 (5.1)	62 (3.9)		149 (7.1)	203 (7.8)		
rs12208240	GG	2800 (79.5)	3354 (79.7)		1231 (86.1)	1350 (84.4)		1569 (75.0)	2004 (76.8)		
	GA	680 (19.3)	807 (19.2)	0.88	187 (13.1)	242 (15.1)	0.32	493 (23.6)	565 (21.7)	0.26	<.0001
	AA	41 (1.2)	49 (1.2)		12 (0.8)	8 (0.5)		29 (1.4)	41 (1.5)		
rs12209785	AA	1850 (52.5)	2211 (52.5)		771 (54.0)	888 (55.5)		1079 (51.6)	1323 (50.7)		
	AG	1413 (40.1)	1720 (40.9)	0.58	563 (39.4)	633 (39.6)	0.14	850 (40.6)	1087 (41.7)	0.75	0.0001
	GG	259 (7.4)	277 (6.6)		95 (6.6)	79 (4.9)		164 (7.8)	198 (7.6)		
rs12333172	CC	2415 (68.6)	2839 (67.4)		932 (65.2)	1021 (63.8)		1483 (70.9)	1818 (69.7)		
	CT	982 (27.9)	1245 (29.6)	0.67	432 (30.2)	514 (32.1)	0.69	550 (26.3)	731 (28.0)	0.66	<.0001
	TT	126 (3.6)	126 (3.0)		66 (4.6)	65 (4.1)		60 (2.9)	61 (2.3)		
rs1316330	GG	2259 (64.1)	2709 (64.4)		783 (54.7)	881 (55.1)		1476 (70.5)	1828 (70.1)		
	GT	1130 (32.1)	1351 (32.1)	0.64	558 (39.0)	634 (39.6)	0.55	572 (27.3)	717 (27.5)	0.64	<.0001
	TT	135 (3.8)	148 (3.5)		90 (6.3)	85 (5.3)		45 (2.2)	63 (2.4)		
rs13201287	GG	1784 (50.6)	2137 (50.8)		768 (53.7)	874 (54.6)		1016 (48.5)	1263 (48.4)		
	GA	1442 (40.9)	1762 (41.9)	0.41	557 (38.9)	642 (40.1)	0.16	885 (42.3)	1120 (42.9)	0.86	<.0001
	AA	298 (8.5)	311 (7.3)		106 (7.4)	84 (5.3)		192 (9.2)	227 (8.7)		
rs1321075	CC	1916 (54.4)	2261 (53.7)	0.61	1020 (71.3)	1137 (71.1)	0.96	896 (42.8)	1124 (43.1)	0.87	<.0001

		Total (n=7,733)			Non-Hispanic White (n=3,030)			Hispanic (n=4,703)			
		Cases (n=3524)	Controls (n=4209)	p[a]	Cases (n=1431)	Controls (n=1599)	p[a]	Cases (n=2093)	Controls (n=2610)	p[a]	p[b]
Gene/SNP	**Genotype[c]**	***N (%)***	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
	CA	1310 (37.2)	1591 (37.8)		369 (25.8)	421 (26.3)		941 (45.0)	1170 (44.8)		
	AA	297 (8.4)	358 (8.5)		42 (2.9)	42 (2.6)		255 (12.2)	316 (12.1)		
rs17209895	TT	2354 (66.8)	2805 (66.6)		769 (53.8)	852 (53.3)		1585 (75.7)	1953 (74.8)		
	TC	1007 (28.6)	1234 (29.3)	0.79	553 (38.7)	643 (40.2)	0.84	454 (21.7)	591 (22.6)	0.56	<.0001
	CC	162 (4.6)	171 (4.1)		108 (7.5)	105 (6.5)		54 (2.6)	66 (2.5)		
rs2396441	CC	911 (25.9)	1034 (24.6)		397 (27.7)	387 (24.2)		514 (24.6)	647 (24.8)		
	CT	1762 (50.0)	2198 (52.2)	0.84	707 (49.4)	832 (52.0)	0.08	1055 (50.4)	1366 (52.4)	0.23	0.30
	TT	851 (24.2)	976 (23.2)		327 (22.9)	380 (23.8)		524 (25.0)	596 (22.8)		
rs2677108	TT	973 (27.6)	1130 (26.9)		499 (34.9)	557 (34.8)		474 (22.6)	573 (22.0)		
	TC	1727 (49.0)	2048 (48.7)	0.26	689 (48.2)	769 (48.1)	0.94	1038 (49.6)	1279 (49.0)	0.36	<.0001
	CC	824 (23.4)	1030 (24.5)		243 (16.9)	274 (17.1)		581 (27.8)	756 (29.0)		
rs2790093	AA	1609 (45.7)	1917 (45.5)		613 (42.8)	701 (43.8)		996 (47.6)	1216 (46.6)		
	AG	1552 (44.0)	1882 (44.7)	0.80	658 (46.0)	737 (46.1)	0.39	894 (42.7)	1145 (43.9)	0.64	0.002
	GG	362 (10.3)	411 (9.8)		160 (11.2)	162 (10.1)		202 (9.7)	249 (9.5)		
rs2819854	TT	898 (25.5)	1113 (26.5)		348 (24.3)	372 (23.3)		550 (26.3)	741 (28.4)		
	TC	1752 (49.7)	2081 (49.5)	0.31	716 (50.0)	818 (51.2)	0.70	1036 (49.5)	1263 (48.4)	0.14	0.0006
	CC	873 (24.8)	1014 (24.1)		367 (25.7)	409 (25.5)		506 (24.2)	605 (23.2)		
rs598953	TT	1266 (35.9)	1480 (35.2)		563 (39.3)	618 (38.6)		703 (33.6)	862 (33.0)		
	TA	1683 (47.8)	2030 (48.2)	0.50	678 (47.4)	771 (48.2)	0.80	1005 (48.0)	1259 (48.2)	0.66	<.0001
	AA	575 (16.3)	700 (16.6)		190 (13.3)	211 (13.2)		385 (18.4)	489 (18.7)		
rs6930053	CC	1518 (43.1)	1821 (43.3)		525 (36.7)	541 (33.8)		993 (47.4)	1280 (49.0)		
	CT	1562 (44.3)	1899 (45.1)	0.47	695 (48.6)	803 (50.2)	0.09	867 (41.4)	1096 (41.9)	0.05	<.0001
	TT	443 (12.6)	490 (11.6)		210 (14.7)	256 (16.0)		233 (11.1)	234 (8.9)		
rs7750470	TT	2252 (63.9)	2718 (64.6)	0.53	886 (61.9)	1040 (65.0)	0.13	1366 (65.3)	1678 (64.3)	0.66	0.48
	TC	1124 (31.9)	1322 (31.4)		488 (34.1)	497 (31.1)		636 (30.4)	825 (31.6)		

		Total (n=7,733)			Non-Hispanic White (n=3,030)			Hispanic (n=4,703)			
		Cases (n=3524)	Controls (n=4209)	p[a]	Cases (n=1431)	Controls (n=1599)	p[a]	Cases (n=2093)	Controls (n=2610)	p[a]	p[b]
Gene/SNP	**Genotype[c]**	***N (%)***	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
	CC	148 (4.2)	170 (4.0)		57 (4.0)	63 (3.9)		91 (4.3)	107 (4.1)		
rs9463090	GG	2281 (64.8)	2743 (65.2)		870 (60.8)	997 (62.4)		1411 (67.5)	1746 (66.9)		
	GA	1075 (30.5)	1288 (30.6)	0.44	474 (33.2)	516 (32.3)	0.28	601 (28.7)	772 (29.6)	0.91	<.0001
	AA	166 (4.7)	175 (4.2)		86 (6.0)	84 (5.3)		80 (3.8)	91 (3.5)		
RUNX3 *(n=8)*											
rs11249206	TT	1247 (36.0)	1472 (35.5)		372 (26.4)	378 (24.0)		875 (42.5)	1094 (42.7)		
	TC	1607 (46.3)	1956 (47.2)	0.97	715 (50.6)	803 (50.9)	0.08	892 (43.4)	1153 (44.9)	0.37	<.0001
	CC	615 (17.7)	715 (17.3)		325 (23.0)	397 (25.1)		290 (14.1)	318 (12.4)		
rs2236850	TT	1189 (33.8)	1486 (35.4)		434 (30.4)	519 (32.5)		755 (36.1)	967 (37.1)		
	TC	1686 (47.9)	1984 (47.2)	0.13	709 (49.6)	764 (47.9)	0.31	977 (46.7)	1220 (46.8)	0.33	<.0001
	CC	645 (18.3)	733 (17.4)		286 (20.0)	312 (19.6)		359 (17.2)	421 (16.1)		
rs4478762	GG	2719 (77.2)	3272 (77.8)		1112 (77.7)	1248 (78.0)		1607 (76.8)	2024 (77.6)		
	GA	751 (21.3)	891 (21.2)	0.31	295 (20.6)	341 (21.3)	0.43	456 (21.8)	550 (21.1)	0.48	0.43
	AA	54 (1.5)	45 (1.0)		24 (1.7)	11 (0.7)		30 (1.4)	34 (1.3)		
rs6688058	GG	2604 (73.9)	3142 (74.6)		1068 (74.6)	1208 (75.5)		1536 (73.4)	1934 (74.1)		
	GA	842 (23.9)	986 (23.4)	0.37	331 (23.1)	368 (23.0)	0.36	511 (24.4)	618 (23.7)	0.64	0.14
	AA	78 (2.2)	82 (2.0)		32 (2.3)	24 (1.5)		46 (2.2)	58 (2.2)		
rs7517302	TT	1260 (35.8)	1537 (36.5)		438 (30.6)	507 (31.7)		822 (39.3)	1030 (39.5)		
	TC	1655 (47.0)	2006 (47.7)	0.17	698 (48.8)	795 (49.7)	0.24	957 (45.8)	1211 (46.5)	0.59	<.0001
	CC	606 (17.2)	664 (15.8)		294 (20.6)	298 (18.6)		312 (14.9)	366 (14.0)		
rs7551188	CC	886 (25.2)	1107 (26.3)		328 (22.9)	343 (21.4)		558 (26.7)	764 (29.3)		
	CT	1756 (49.9)	2037 (48.5)	0.59	714 (49.9)	792 (49.5)	0.19	1042 (49.9)	1245 (47.8)	0.13	<.0001
	TT	878 (24.9)	1061 (25.2)		389 (27.2)	465 (29.1)		489 (23.4)	596 (22.9)		
rs906296	CC	2143 (60.9)	2701 (64.2)	0.005	820 (57.3)	952 (59.5)	0.38	1323 (63.2)	1749 (67.1)	0.007	<.0001
	CG	1208 (34.3)	1315 (31.3)		533 (37.3)	558 (34.9)		675 (32.3)	758 (29.1)		

		Total (n=7,733)			Non-Hispanic White (n=3,030)			Hispanic (n=4,703)			
		Cases (n=3524)	Controls (n=4209)	p[a]	Cases (n=1431)	Controls (n=1599)	p[a]	Cases (n=2093)	Controls (n=2610)	p[a]	p[b]
Gene/SNP	**Genotype[c]**	***N (%)***	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
	GG	171 (4.9)	190 (4.5)		77 (5.4)	90 (5.6)		94 (4.5)	100 (3.8)		
rs9438876	AA	1016 (28.9)	1265 (30.0)		308 (21.5)	366 (22.9)		708 (33.8)	899 (34.4)		
	AG	1713 (48.6)	2005 (47.6)	0.39	713 (49.9)	751 (46.9)	0.92	1000 (47.8)	1254 (48.1)	0.48	<.0001
	GG	793 (22.5)	940 (22.3)		409 (28.6)	483 (30.2)		384 (18.4)	457 (17.5)		
TGFB1 ***(n=2)***											
rs1800469	CC	1253 (36.2)	1523 (36.6)		664 (47.6)	753 (47.9)		589 (28.5)	770 (29.7)		
	CT	1590 (45.9)	1913 (46.0)	0.57	585 (41.9)	642 (40.8)	0.87	1005 (48.7)	1271 (49.2)	0.15	<.0001
	TT	618 (17.9)	722 (17.4)		147 (10.5)	177 (11.3)		471 (22.8)	545 (21.1)		
rs4803455	CC	1193 (37.0)	1401 (35.4)		368 (26.2)	409 (26.0)		825 (45.3)	992 (41.6)		
	CA	1460 (45.3)	1890 (47.7)	0.62	698 (49.7)	780 (49.5)	0.80	762 (41.9)	1110 (46.5)	0.18	<.0001
	AA	572 (17.7)	671 (16.9)		339 (24.1)	387 (24.5)		233 (12.8)	284 (11.9)		
TGFB-R1 ***(n=5)***											
rs1013186	GG	2578 (73.2)	3109 (73.9)		912 (63.7)	1026 (64.1)		1666 (79.6)	2083 (79.8)		
	GA	874 (24.8)	1016 (24.1)	0.53	475 (33.2)	513 (32.1)	0.86	399 (19.1)	503 (19.3)	0.62	<.0001
	AA	72 (2.0)	85 (2.0)		44 (3.1)	61 (3.8)		28 (1.3)	24 (0.9)		
rs10733710	GG	1834 (52.1)	2092 (49.7)		906 (63.4)	954 (59.7)		928 (44.3)	1138 (43.6)		
	GA	1387 (39.4)	1694 (40.3)	0.01	467 (32.7)	565 (35.3)	0.03	920 (44.0)	1129 (43.3)	0.28	<.0001
	AA	302 (8.5)	423 (10.0)		57 (3.9)	80 (5.0)		245 (11.7)	343 (13.1)		
rs11568785	AA	3093 (87.8)	3730 (88.6)		1172 (81.9)	1343 (83.9)		1921 (91.8)	2387 (91.5)		
	AG	410 (11.6)	463 (11.0)	0.19	245 (17.1)	243 (15.2)	0.15	165 (7.9)	220 (8.4)	0.90	<.0001
	GG	21 (0.6)	17 (0.4)		14 (1.0)	14 (0.9)		7 (0.3)	3 (0.1)		
rs1571590	AA	2580 (73.2)	3112 (73.9)		912 (63.7)	1026 (64.2)		1668 (79.7)	2086 (79.9)		
	AG	872 (24.7)	1014 (24.1)	0.48	475 (33.2)	512 (32.0)	0.88	397 (19.0)	502 (19.2)	0.57	<.0001
	GG	72 (2.0)	83 (1.9)		44 (3.1)	61 (3.8)		28 (1.3)	22 (0.8)		
rs6478974	TT	1247 (35.4)	1574 (37.4)	0.01	398 (27.8)	471 (29.4)	0.26	849 (40.6)	1103 (42.3)	0.06	<.0001

		Total (n=7,733)			Non-Hispanic White (n=3,030)			Hispanic (n=4,703)			
		Cases (n=3524)	Controls (n=4209)	p[a]	Cases (n=1431)	Controls (n=1599)	p[a]	Cases (n=2093)	Controls (n=2610)	p[a]	p[b]
Gene/SNP	**Genotype[c]**	***N (%)***	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
	TA	1637 (46.5)	1958 (46.5)		690 (48.3)	768 (48.0)		947 (45.2)	1190 (45.6)		
	AA	639 (18.1)	677 (16.1)		342 (23.9)	361 (22.6)		297 (14.2)	316 (12.1)		

Note: Percentages are rounded and may not total to 100%

[a] Case-control comparison for entire study population and within each ethnic group (p-values reported from *Mantel-Haenszel* chi-square tests)

[b] Ethnic group comparison, regardless of case-control status (*p*-values reported from *Mantel-Haenszel* chi-square tests)

[c] In vertical order for each SNP: homozygous wild-type, heterozygote, homozygous variant

Table 8. ERα: genotype distribution, New Mexico subpopulation by self-reported race and case-control status

		Study Population			Study Population Stratified by Self-Reported Race						
Gene/SNP	**Genotype[c]**	**Total (n=1,409)**			**Non-Hispanic White (n=927)**			**Hispanic (n=482)**			
		Cases (n=694)	Controls (n=715)	p[a]	Cases (n=452)	Controls (n=475)	p[a]	Cases (n=242)	Controls (n=240)	p[a]	p[b]
		N (%)	***N (%)***		***N (%)***	***N (%)***		***N (%)***	***N (%)***		
ERα *(n=5)*											
rs1801132	CC	383 (55.2)	415 (58.0)		252 (55.6)	286 (60.2)		131 (54.1)	129 (53.8)		
	CG	258 (37.2)	267 (37.3)	0.07	171 (37.8)	171 (36.0)	0.07	87 (36.0)	96 (40.0)	0.57	0.04
	GG	53 (7.6)	33 (4.6)		29 (6.4)	18 (3.8)		24 (9.9)	15 (6.2)		
rs3798577	TT	215 (31.0)	222 (31.1)		128 (28.3)	138 (29.1)		87 (36.0)	84 (35.0)		
	TC	318 (45.8)	362 (50.6)	0.19	210 (46.5)	244 (51.4)	0.17	108 (44.6)	118 (49.2)	0.68	0.004
	CC	161 (23.2)	131 (18.3)		114 (25.2)	93 (19.6)		47 (19.4)	38 (15.8)		
rs2046210	GG	288 (41.5)	328 (45.9)		169 (37.4)	198 (41.7)		119 (49.2)	130 (54.2)		
	GA	322 (46.4)	298 (41.7)	0.27	223 (49.3)	211 (44.4)	0.41	99 (40.9)	87 (36.3)	0.38	<.0001
	AA	84 (12.1)	89 (12.5)		60 (13.3)	66 (13.9)		24 (9.9)	23 (9.6)		
rs851984	GG	255 (36.7)	274 (38.3)		167 (36.9)	186 (39.2)		88 (36.4)	88 (36.7)		
	GA	327 (47.1)	336 (47.0)	0.41	213 (47.1)	221 (46.5)	0.40	114 (47.1)	115 (47.9)	0.82	0.53
	AA	112 (16.1)	105 (14.7)		72 (15.9)	68 (14.3)		40 (16.5)	37 (15.4)		
rs6913578	AA	315 (45.4)	352 (49.2)		190 (42.0)	215 (45.3)		125 (51.7)	137 (57.1)		
	AC	302 (43.5)	293 (41.0)	0.15	208 (46.0)	208 (43.8)	0.34	94 (38.8)	85 (35.4)	0.21	0.0003
	CC	77 (11.1)	70 (9.8)		54 (12.0)	52 (10.9)		23 (9.5)	18 (7.5)		

Note: Percentages are rounded and may not total to 100%

[a] Case-control comparison for entire study population and within each ethnic group (p-values reported from *Mantel-Haenszel* chi-square tests

[b] Ethnic group comparison, regardless of case-control status (*p*-values reported from *Mantel-Haenszel* chi-square tests)

[c] In vertical order for each SNP: homozygous wild-type, heterozygote, homozygous variant

Table 11b: TGF-β signaling and ERα genes: overall associations with breast cancer risk: *The Breast Cancer Health Disparities Study (Full table)*

	Controls		Cases					
	N	(%)	N	(%)	OR[a]	(95% CI)		p[b]
RUNX1 (rs7279383)								
CC	3098	73.6	2647	75.1	1.00			0.03 (0.23)
CG/GG	1110	26.4	876	24.9	0.89	(0.80,	0.99)	
RUNX1 (rs2268288)								
TT/TC	4107	97.6	3415	96.9	1.00			0.14 (0.75)
CC	102	2.4	108	3.1	1.23	(0.93,	1.62)	
RUNX1 (rs2252585)								
TT/TC	3585	85.2	3040	86.3	1.00			0.45 (1.00))
CC	624	14.8	483	13.7	0.95	(0.83,	1.08)	
RUNX1 (rs11701453)								
GG	2917	69.3	2395	68.0	1.00			
GC	1168	27.8	1030	29.2	1.06	(0.96,	1.17)	0.59 (1.00)
CC	124	2.9	97	2.8	0.91	(0.70,	1.20)	
RUNX1 (rs8127225)								
TT	2591	61.6	2117	60.1	1.00			0.03 (0.23)
TC/CC	1612	38.4	1407	39.9	1.11	(1.01,	1.22)	
RUNX1 (rs1474479)								
GG/GA	3874	92.0	3251	92.3	1.00			0.33 (1.00)
AA	335	8.0	272	7.7	0.92	(0.77,	1.09)	
RUNX1 (rs1883066)								
GG	3470	82.5	2856	81.0	1.00			0.29 (1.00)
GC/CC	738	17.5	668	19.0	1.07	(0.95,	1.20)	
RUNX1 (rs7279123)								
CC	2628	62.6	2236	63.7	1.00			
CT	1361	32.4	1109	31.6	0.94	(0.85,	1.03)	0.13 (0.75)
TT	206	4.9	163	4.6	0.90	(0.72,	1.11)	
RUNX2 (rs1321075)								
CC	2260	53.7	1916	54.4	1.00			
CA	1591	37.8	1310	37.2	1.01	(0.92,	1.12)	0.56 (1.00)
AA	358	8.5	297	8.4	1.06	(0.89,	1.26)	
RUNX2 (rs17209895)								
TT	2804	66.6	2354	66.8	1.00			
TC	1234	29.3	1007	28.6	0.93	(0.84,	1.03)	0.51 (1.00)
CC	171	4.1	162	4.6	1.06	(0.84,	1.32)	
RUNX2 (rs2677108)								
TT	1129	26.8	973	27.6	1.00			
TC	2048	48.7	1727	49.0	0.99	(0.89,	1.11)	0.55 (1.00)
CC	1030	24.5	824	23.4	0.96	(0.85,	1.09)	
RUNX2 (rs2819854)								
TT	1113	26.5	898	25.5	1.00			
TC	2080	49.4	1752	49.7	1.04	(0.93,	1.16)	0.36 (1.00)
CC	1014	24.1	873	24.8	1.06	(0.94,	1.20)	

	Controls		Cases					
	N	(%)	N	(%)	OR[a]	(95% CI)		p[b]
RUNX2 (rs2790093)								
AA	1917	45.5	1609	45.7	1.00			
AG	1882	44.7	1552	44.1	0.98	(0.89,	1.08)	0.80 (1.00)
GG	410	9.7	362	10.3	1.05	(0.90,	1.23)	
RUNX2 (rs9463090)								
GG	2742	65.2	2281	64.8	1.00			
GA	1288	30.6	1075	30.5	0.99	(0.90,	1.10)	0.60 (1.00)
AA	175	4.2	166	4.7	1.12	(0.89,	1.39)	
RUNX2 (rs2396441)								
CC	1034	24.6	911	25.9	1.00			
CT	2198	52.2	1762	50.0	0.91	(0.82,	1.02)	0.81 (1.00)
TT	976	23.2	851	24.1	0.99	(0.87,	1.12)	
RUNX2 (rs1316330)								
GG	2709	64.4	2259	64.1	1.00			
GT	1350	32.1	1130	32.1	0.98	(0.89,	1.08)	0.89 (1.00)
TT	148	3.5	135	3.8	1.04	(0.82,	1.33)	
RUNX2 (rs7750470)								
TT	2717	64.6	2252	63.9	1.00			
TC	1322	31.4	1124	31.9	1.02	(0.93,	1.13)	0.56 (1.00)
CC	170	4.0	148	4.2	1.05	(0.84,	1.32)	
RUNX2 (rs6930053)								
CC	1820	43.2	1518	43.1	1.00			
CT	1899	45.1	1562	44.3	0.96	(0.87,	1.06)	0.94 (1.00)
TT	490	11.6	443	12.6	1.04	(0.90,	1.21)	
RUNX2 (rs12208240)								
GG	3353	79.7	2800	79.5	1.00			
GA	807	19.2	680	19.3	1.03	(0.92,	1.16)	0.58 (1.00)
AA	49	1.2	41	1.2	1.04	(0.69,	1.59)	
RUNX2 (rs12209785)								
AA/AG	3931	93.4	3263	92.6	1.00			0.13 (1.00)
GG	276	6.6	259	7.4	1.15	(0.96,	1.37)	
RUNX2 (rs10948238)								
CC/CT	3602	85.6	2950	83.8	1.00			0.03 (0.42)
TT	604	14.4	571	16.2	1.15	(1.01,	1.30)	
RUNX2 (rs13201287)								
GG/GA	3899	92.6	3226	91.5	1.00			0.05 (0.69)
AA	310	7.4	298	8.5	1.18	(1.00,	1.39)	
RUNX2 (rs12333172)								
CC/CT	4083	97.0	3397	96.4	1.00			0.26 (1.00)
TT	126	3.0	126	3.6	1.16	(0.90,	1.49)	

	Controls		Cases					
	N	(%)	N	(%)	OR[a]	(95% CI)		p[b]
RUNX2 (rs1200428)								
CC	2316	55.0	1932	54.8	1.00			
CA	1628	38.7	1370	38.9	1.02	(0.93,	1.13)	0.61 (1.00)
AA	265	6.3	222	6.3	1.03	(0.85,	1.24)	
RUNX2 (rs598953)								
TT	1480	35.2	1266	35.9	1.00			
TA	2029	48.2	1683	47.8	0.98	(0.89,	1.09)	0.76 (1.00)
AA	700	16.6	575	16.3	0.98	(0.86,	1.12)	
RUNX3 (rs2236850)								
TT	1486	35.4	1189	33.8	1.00			
TC	1983	47.2	1686	47.9	1.05	(0.95,	1.16)	0.18 (0.75)
CC	733	17.4	645	18.3	1.09	(0.95,	1.24)	
RUNX3 (rs9438876)								
AA	1265	30.1	1016	28.8	1.00			
AG	2004	47.6	1713	48.6	1.04	(0.94,	1.16)	0.75 (1.00)
GG	940	22.3	793	22.5	1.02	(0.89,	1.16)	
RUNX3 (rs7517302)								
TT/TC	3542	84.2	2915	82.8	1.00			0.12 (0.60)
CC	664	15.8	606	17.2	1.10	(0.98,	1.24)	
RUNX3 (rs906296)								
CC	2701	64.2	2143	60.8	1.00			0.004 (0.03)
CG/GG	1505	35.8	1379	39.2	1.15	(1.04,	1.26)	
RUNX3 (rs7551188)								
CC	1106	26.3	886	25.2	1.00			
CT	2037	48.5	1756	49.9	1.06	(0.95,	1.18)	0.90 (1.00)
TT	1061	25.2	878	24.9	1.01	(0.89,	1.14)	
RUNX3 (rs6688058)								
GG/GA	4127	98.1	3446	97.8	1.00			0.36 (1.00)
AA	82	1.9	78	2.2	1.16	(0.85,	1.58)	
RUNX3 (rs11249206)								
TT	1471	35.5	1247	35.9	1.00			
TC	1956	47.2	1607	46.3	0.94	(0.84,	1.04)	0.38 (1.00)
CC	715	17.3	615	17.7	0.96	(0.83,	1.10)	
RUNX3 (rs4478762)								
GG/GA	4162	98.9	3470	98.5	1.00			0.07 (0.41)
AA	45	1.1	54	1.5	1.45	(0.97,	2.17)	
TGF-β1 (rs1800469)								
CC/CT	3435	82.6	2843	82.1	1.00			0.21 (0.21)
TT	722	17.4	618	17.9	1.08	(0.96,	1.22)	
TGF-β1 (rs4803455)								
CC	1400	35.3	1193	37.0	1.00			0.02 (0.04)
CA/AA	2561	64.7	2032	63.0	0.89	(0.81,	0.98)	

	Controls		Cases					
	N	(%)	N	(%)	OR[a]	(95% CI)		p[b]
TGF- βR1 (rs6478974)								
TT/TA	3531	83.9	2884	81.9	1.00			0.05 (0.19)
AA	677	16.1	639	18.1	1.13	(1.00,	1.28)	
TGF- βR1 (rs1571590)								
AA	3111	73.9	2580	73.2	1.00			0.89 (1.00)
AG	1014	24.1	872	24.7	0.99	(0.89,	1.11)	
GG	83	2.0	72	2.0	0.98	(0.71,	1.35)	
TGF- βR1 (rs1013186)								
GG	3108	73.8	2578	73.2	1.00			0.85 (1.00)
GA	1016	24.1	874	24.8	1.00	(0.89,	1.11)	
AA	85	2.0	72	2.0	0.96	(0.70,	1.32)	
TGF- βR1 (rs11568785)								
AA/AG	4192	99.6	3503	99.4	1.00			0.29 (0.67)
GG	17	0.4	21	0.6	1.41	(0.74,	2.69)	
TGF- βR1 (rs10733710)								
GG/GA	3785	89.9	3221	91.4	1.00			0.12 (0.40)
AA	423	10.1	302	8.6	0.88	(0.75,	1.03)	
ERα (rs1801132)								
CC/CG	672	95.3	630	92.2	1.00			0.02 (0.08)
GG	33	4.7	53	7.8	1.72	(1.10,	2.69)	
ERα (rs3798577)								
TT/TC	577	81.8	526	77.0	1.00			0.02 (0.08)
CC	128	18.2	157	23.0	1.36	(1.04,	1.76)	
ERα (rs2046210)								
GG	322	45.7	281	41.1	1.00			0.07 (0.19)
GA/AA	383	54.3	402	58.9	1.22	(0.98,	1.51)	
ERα (rs851984)								
GG	272	38.6	251	36.8	1.00			0.37 (0.37)
GA	330	46.8	322	47.1	1.06	(0.84,	1.33)	
AA	103	14.6	110	16.1	1.16	(0.84,	1.59)	
ERα (rs6913578)								
AA	346	49.1	308	45.1	1.00			0.11 (0.19)
AC	289	41.0	298	43.6	1.17	(0.93,	1.46)	
CC	70	9.9	77	11.3	1.26	(0.88,	1.81)	

[a] TGF-β signaling SNPs (n=7733) adjusted for age, study site, and genetic admixture; ERα SNPs adjusted for age and genetic admixture (n=1388)

[b] Wald p-value for 1 df test; Bonferroni-Holm p-value for adjustment for multiple comparisons shown in parentheses

Table 12b. The association of TGF-β signaling and ERα genes and breast cancer stratified by menopausal status *(Full table)*

	Pre/Peri-Menopause							Post-Menopause							
	Controls		Cases					Controls		Cases					
	N	(%)	N	(%)	OR [a]	(95% CI)		N	(%)	N	(%)	OR [a]	(95% CI)		p-int [b]
RUNX1 (rs7279383)															
CC	1131	74.4	980	75.0	1.00			1884	73.2	1583	75.2	1.00			0.67 (1.00)
CG/GG	390	25.6	326	25.0	0.91	(0.77,	1.09)	690	26.8	522	24.8	0.88	(0.77,	1.01)	
Wald-p [c]					1.00							0.39			
RUNX1 (rs2268288)															
TT/TC	1480	97.3	1273	97.5	1.00			2515	97.7	2033	96.6	1.00			0.10 (0.77)
CC	41	2.7	33	2.5	0.88	(0.55,	1.41)	59	2.3	71	3.4	1.47	(1.03,	2.09)	
Wald-p [c]					1.00							0.26			
RUNX1 (rs2252585)															
TT/TC	1268	83.4	1098	84.1	1.00			2214	86.0	1847	87.8	1.00			0.41 (1.00)
CC	253	16.6	208	15.9	1.01	(0.82,	1.24)	360	14.0	257	12.2	0.88	(0.74,	1.05)	
Wald-p [c]					1.00							0.68			
RUNX1 (rs11701453)															
GG	1084	71.3	911	69.8	1.00			1755	68.2	1414	67.2	1.00			0.26 (1.00)
GC	398	26.2	352	27.0	1.03	(0.87,	1.22)	739	28.7	638	30.3	1.07	(0.94,	1.21)	
CC	39	2.6	42	3.2	1.23	(0.79,	1.93)	80	3.1	52	2.5	0.78	(0.54,	1.11)	
p-trend [c]					1.00							1.00			
RUNX1 (rs8127225)															
TT	919	60.6	747	57.2	1.00			1602	62.3	1303	61.9	1.00			0.15 (0.92)
TC/CC	598	39.4	559	42.8	1.24	(1.06,	1.44)	970	37.7	802	38.1	1.04	(0.92,	1.17)	
Wald-p [c]					0.06							1.00			
RUNX1 (rs1474479)															
GG/GA	1422	93.5	1207	92.5	1.00			2344	91.1	1943	92.3	1.00			0.10 (0.77)
AA	99	6.5	98	7.5	1.08	(0.80,	1.46)	230	8.9	162	7.7	0.81	(0.65,	1.01)	
Wald-p [c]					1.00							0.39			

	Pre/Peri-Menopause							Post-Menopause							
	Controls		Cases					Controls		Cases					
	N	(%)	N	(%)	OR [a]	(95% CI)		N	(%)	N	(%)	OR [a]	(95% CI)		p-int [b]
RUNX1 (rs1883066)															
GG	1279	84.1	1071	82.0	1.00			2092	81.3	1694	80.5	1.00			0.52 (1.00)
GC/CC	242	15.9	235	18.0	1.11	(0.91,	1.35)	481	18.7	411	19.5	1.03	(0.89,	1.20)	
Wald-p [c]					1.00							1.00			
RUNX1 (rs7279123)															
CC	972	64.1	835	64.2	1.00			1583	61.7	1326	63.3	1.00			0.26 (1.00)
CT	486	32.0	402	30.9	0.94	(0.79,	1.10)	841	32.8	671	32.0	0.93	(0.82,	1.06)	
TT	59	3.9	63	4.8	1.19	(0.82,	1.72)	141	5.5	99	4.7	0.81	(0.62,	1.07)	
p-trend [c]					1.00							0.48			
RUNX2 (rs1321075)															
CC	770	50.6	706	54.1	1.00			1427	55.4	1147	54.5	1.00			0.24 (1.00)
CA	604	39.7	483	37.0	0.92	(0.78,	1.08)	944	36.7	794	37.7	1.08	(0.95,	1.23)	
AA	147	9.7	116	8.9	0.94	(0.71,	1.24)	203	7.9	164	7.8	1.09	(0.87,	1.37)	
p-trend [c]					1.00							1.00			
RUNX2 (rs17209895)															
TT	1043	68.6	885	67.8	1.00			1685	65.5	1391	66.1	1.00			0.82 (1.00)
TC	422	27.7	368	28.2	0.96	(0.81,	1.14)	775	30.1	610	29.0	0.92	(0.81,	1.05)	
CC	56	3.7	52	4.0	0.97	(0.65,	1.44)	114	4.4	104	4.9	1.06	(0.81,	1.41)	
p-trend [c]					1.00							1.00			
RUNX2 (rs2677108)															
TT	411	27.0	369	28.3	1.00			686	26.7	570	27.1	1.00			0.37 (1.00)
TC	709	46.6	635	48.6	1.01	(0.85,	1.21)	1283	49.9	1044	49.6	0.99	(0.86,	1.14)	
CC	400	26.3	302	23.1	0.88	(0.71,	1.08)	604	23.5	491	23.3	1.01	(0.85,	1.19)	
p-trend [c]					1.00							1.00			
RUNX2 (rs2819854)															
TT	392	25.8	319	24.4	1.00			689	26.8	545	25.9	1.00			0.90 (1.00)

	Pre/Peri-Menopause						Post-Menopause						
	Controls		Cases				Controls		Cases				
	N	(%)	N	(%)	OR [a]	(95% CI)	N	(%)	N	(%)	OR [a]	(95% CI)	p-int [b]
TC	747	49.1	641	49.1	1.05	(0.87, 1.26)	1277	49.7	1058	50.3	1.04	(0.91, 1.20)	
CC	382	25.1	346	26.5	1.11	(0.90, 1.36)	606	23.6	501	23.8	1.04	(0.89, 1.23)	
p-trend [c]					1.00						1.00		
RUNX2 (rs2790093)													
AA	693	45.6	626	48.0	1.00		1177	45.7	938	44.6	1.00		0.22 (1.00)
AG	677	44.5	559	42.8	0.91	(0.78, 1.06)	1149	44.6	945	44.9	1.03	(0.92, 1.17)	
GG	151	9.9	120	9.2	0.87	(0.66, 1.13)	248	9.6	222	10.5	1.13	(0.92, 1.38)	
p-trend [c]					1.00						1.00		
RUNX2 (rs9463090)													
GG	991	65.2	850	65.1	1.00		1676	65.2	1361	64.7	1.00		0.13 (1.00)
GA	475	31.3	389	29.8	0.94	(0.80, 1.11)	775	30.1	649	30.8	1.03	(0.90, 1.17)	
AA	54	3.6	67	5.1	1.42	(0.98, 2.06)	120	4.7	94	4.5	0.94	(0.71, 1.24)	
p-trend [c]					1.00						1.00		
RUNX2 (rs2396441)													
CC	369	24.3	330	25.3	1.00		635	24.7	538	25.6	1.00		0.12 (1.00)
CT	808	53.1	638	48.9	0.88	(0.74, 1.06)	1326	51.5	1081	51.4	0.96	(0.84, 1.11)	
TT	344	22.6	338	25.9	1.10	(0.89, 1.36)	612	23.8	486	23.1	0.94	(0.79, 1.10)	
p-trend [c]					1.00						1.00		
RUNX2 (rs1316330)													
GG	984	64.7	853	65.3	1.00		1657	64.4	1334	63.4	1.00		0.48 (1.00)
GT	483	31.8	411	31.5	0.94	(0.80, 1.11)	827	32.1	685	32.5	1.01	(0.89, 1.15)	
TT	53	3.5	42	3.2	0.84	(0.55, 1.27)	89	3.5	86	4.1	1.16	(0.85, 1.58)	
p-trend [c]					1.00						1.00		
RUNX2 (rs7750470)													
TT	986	64.8	824	63.1	1.00		1662	64.6	1360	64.6	1.00		0.62 (1.00)
TC	472	31.0	427	32.7	1.08	(0.92, 1.27)	815	31.7	658	31.3	0.98	(0.87, 1.12)	

	Pre/Peri-Menopause							Post-Menopause							
	Controls		Cases					Controls		Cases					
	N	(%)	N	(%)	OR [a]	(95% CI)		N	(%)	N	(%)	OR [a]	(95% CI)		p-int [b]
CC	63	4.1	55	4.2	1.05	(0.72,	1.52)	97	3.8	87	4.1	1.09	(0.81,	1.47)	
p-trend [c]					1.00							1.00			
RUNX2 (rs6930053)															
CC	686	45.1	570	43.7	1.00			1086	42.2	890	42.3	1.00			0.17 (1.00)
CT	678	44.6	565	43.3	0.97	(0.83,	1.14)	1167	45.3	958	45.5	0.98	(0.87,	1.11)	
TT	157	10.3	170	13.0	1.24	(0.97,	1.58)	321	12.5	257	12.2	0.94	(0.78,	1.14)	
p-trend [c]					1.00							1.00			
RUNX2 (rs12208240)															
GG	1198	78.8	1008	77.2	1.00			2070	80.4	1696	80.7	1.00			0.30 (1.00)
GA	301	19.8	284	21.7	1.16	(0.96,	1.39)	478	18.6	381	18.1	0.99	(0.85,	1.15)	
AA	22	1.4	14	1.1	0.80	(0.40,	1.58)	26	1.0	25	1.2	1.20	(0.69,	2.09)	
p-trend [c]					1.00							1.00			
RUNX2 (rs12209785)															
AA/AG	1416	93.2	1217	93.2	1.00			2408	93.6	1948	92.6	1.00			0.37 (1.00)
GG	104	6.8	89	6.8	1.00	(0.74,	1.34)	165	6.4	155	7.4	1.18	(0.94,	1.49)	
Wald-p [c]					1.00							1.00			
RUNX2 (rs10948238)															
CC/CT	1306	86.0	1095	83.8	1.00			2197	85.4	1771	84.2	1.00			0.58 (1.00)
TT	213	14.0	211	16.2	1.17	(0.95,	1.44)	376	14.6	332	15.8	1.09	(0.93,	1.28)	
Wald-p [c]					1.00							1.00			
RUNX2 (rs13201287)															
GG/GA	1406	92.4	1201	92.0	1.00			2386	92.7	1925	91.4	1.00			0.54 (1.00)
AA	115	7.6	105	8.0	1.08	(0.82,	1.42)	188	7.3	180	8.6	1.20	(0.97,	1.49)	
Wald-p [c]					1.00							0.93			
RUNX2 (rs12333172)															
CC/CT	1479	97.2	1261	96.6	1.00			2492	96.8	2025	96.2	1.00			0.84 (1.00)

	Pre/Peri-Menopause						Post-Menopause						
	Controls		Cases				Controls		Cases				
	N	(%)	N	(%)	OR [a]	(95% CI)	N	(%)	N	(%)	OR [a]	(95% CI)	p-int [b]
TT	42	2.8	45	3.4	1.22	(0.79, 1.87)	82	3.2	79	3.8	1.14	(0.83, 1.57)	
Wald-p [c]					1.00						1.00		
RUNX2 (rs1200428)													
CC	829	54.5	726	55.6	1.00		1422	55.2	1149	54.6	1.00		0.45 (1.00)
CA	591	38.9	506	38.7	0.99	(0.85, 1.16)	995	38.7	819	38.9	1.04	(0.92, 1.17)	
AA	101	6.6	74	5.7	0.86	(0.63, 1.18)	157	6.1	137	6.5	1.10	(0.86, 1.41)	
p-trend [c]					1.00						1.00		
RUNX2 (rs598953)													
TT	531	34.9	476	36.4	1.00		903	35.1	752	35.7	1.00		0.41 (1.00)
TA	734	48.3	634	48.5	0.97	(0.82, 1.14)	1245	48.4	995	47.3	0.97	(0.85, 1.11)	
AA	256	16.8	196	15.0	0.87	(0.70, 1.09)	426	16.6	358	17.0	1.03	(0.87, 1.23)	
p-trend [c]					1.00						1.00		
RUNX3 (rs2236850)													
TT	579	38.1	451	34.5	1.00		875	34.1	715	34.0	1.00		0.30 (1.00)
TC	698	45.9	630	48.2	1.15	(0.97, 1.35)	1222	47.6	993	47.3	0.98	(0.86, 1.12)	
CC	243	16.0	225	17.2	1.17	(0.94, 1.46)	471	18.3	393	18.7	1.01	(0.86, 1.20)	
p-trend [c]					0.41						1.00		
RUNX3 (rs9438876)													
AA	484	31.8	385	29.5	1.00		756	29.4	602	28.6	1.00		0.74 (1.00)
AG	726	47.7	644	49.3	1.09	(0.91, 1.29)	1214	47.2	1016	48.3	1.03	(0.90, 1.19)	
GG	311	20.4	277	21.2	1.07	(0.86, 1.33)	604	23.5	485	23.1	0.98	(0.83, 1.16)	
p-trend [c]					0.47						1.00		
RUNX3 (rs7517302)													
TT/TC	1299	85.5	1091	83.5	1.00		2146	83.4	1736	82.6	1.00		0.52 (1.00)
CC	221	14.5	215	16.5	1.14	(0.93, 1.40)	426	16.6	366	17.4	1.06	(0.91, 1.24)	
Wald-p [c]					0.44						1.00		

	Pre/Peri-Menopause						Post-Menopause						
	Controls		Cases				Controls		Cases				
	N	(%)	N	(%)	OR [a]	(95% CI)	N	(%)	N	(%)	OR [a]	(95% CI)	p-int [b]
RUNX3 (rs906296)													
CC	1009	66.3	774	59.3	1.00		1619	63.0	1301	61.9	1.00		0.01 (0.08)
CG/GG	512	33.7	532	40.7	**1.33**	**(1.14, 1.55)**	952	37.0	802	38.1	1.05	(0.93, 1.18)	
Wald-p [c]					**0.002**						1.00		
RUNX3 (rs7551188)													
CC	392	25.8	351	26.9	1.00		683	26.6	516	24.5	1.00		0.23 (1.00)
CT	753	49.6	659	50.5	0.96	(0.80, 1.15)	1236	48.1	1047	49.8	1.11	(0.96, 1.28)	
TT	374	24.6	296	22.7	0.85	(0.68, 1.05)	653	25.4	540	25.7	1.08	(0.92, 1.27)	
p-trend [c]					0.41						1.00		
RUNX3 (rs6688058)													
GG/GA	1492	98.1	1284	98.3	1.00		2526	98.1	2053	97.5	1.00		0.32 (1.00)
AA	29	1.9	22	1.7	0.93	(0.53, 1.63)	48	1.9	52	2.5	1.33	(0.89, 1.98)	
Wald-p [c]					0.80						0.82		
RUNX3 (rs11249206)													
TT	550	36.8	499	38.7	1.00		888	35.0	714	34.5	1.00		0.25 (1.00)
TC	695	46.5	587	45.5	0.88	(0.74, 1.04)	1206	47.6	969	46.8	0.98	(0.85, 1.11)	
CC	251	16.8	203	15.7	0.80	(0.64, 1.01)	442	17.4	389	18.8	1.06	(0.89, 1.26)	
p-trend [c]					0.23						1.00		
RUNX3 (rs4478762)													
GG/GA	1504	98.9	1293	99.0	1.00		2546	99.0	2068	98.2	1.00		0.21 (1.00)
AA	16	1.1	13	1.0	0.98	(0.47, 2.05)	27	1.0	37	1.8	1.71	(1.04, 2.82)	
Wald-p [c]					0.95						0.21		
TGF-β1 (rs1800469)													
CC	522	34.9	458	35.6	1.00		963	37.8	753	36.5	1.00		0.53 (1.00)
CT	716	47.8	591	46.0	0.99	(0.84, 1.17)	1139	44.7	947	45.9	1.08	(0.95, 1.23)	
TT	259	17.3	237	18.4	1.14	(0.91, 1.42)	445	17.5	362	17.6	1.09	(0.91, 1.29)	

	Pre/Peri-Menopause						Post-Menopause						
	Controls		Cases				Controls		Cases				
	N	(%)	N	(%)	OR [a]	(95% CI)	N	(%)	N	(%)	OR [a]	(95% CI)	p-int [b]
p-trend [c]					0.34						0.28		
TGF-β1 (rs4803455)													
CC	510	36.3	462	38.9	1.00		861	35.0	709	36.4	1.00		0.52 (1.00)
CA/AA	896	63.7	727	61.1	0.84	(0.71, 0.99)	1598	65.0	1238	63.6	0.91	(0.80, 1.03)	
Wald-p [c]					0.07						0.28		
TGF- βRI (rs6478974)													
TT/TA	1273	83.8	1077	82.5	1.00		2156	83.8	1715	81.5	1.00		0.51 (1.00)
AA	247	16.3	229	17.5	1.04	(0.85, 1.27)	418	16.2	389	18.5	1.16	(1.00, 1.35)	
Wald-p [c]					0.77						0.23		
TGF- βRI (rs1571590)													
AA	1155	75.9	962	73.7	1.00		1872	72.8	1535	72.9	1.00		0.55 (1.00)
AG	342	22.5	317	24.3	1.06	(0.89, 1.27)	648	25.2	526	25.0	0.96	(0.83, 1.10)	
GG	24	1.6	27	2.1	1.20	(0.68, 2.11)	53	2.1	44	2.1	0.95	(0.63, 1.43)	
p-trend [c]					0.77						1.00		
TGF- βRI (rs1013186)													
GG	1156	76.0	960	73.5	1.00		1869	72.6	1535	72.9	1.00		0.49 (1.00)
GA	339	22.3	319	24.4	1.08	(0.90, 1.29)	652	25.3	526	25.0	0.95	(0.83, 1.09)	
AA	26	1.7	27	2.1	1.12	(0.64, 1.94)	53	2.1	44	2.1	0.95	(0.63, 1.42)	
p-trend [c]					0.77						1.00		
TGF- βRI (rs11568785)													
AA/AG	1517	99.7	1298	99.4	1.00		2562	99.5	2092	99.4	1.00		0.52 (1.00)
GG	4	0.3	8	0.6	2.01	(0.60, 6.71)	12	0.5	13	0.6	1.32	(0.60, 2.91)	
Wald-p [c]					0.77						1.00		
TGF- βRI (rs10733710)													
GG/GA	1356	89.2	1202	92.0	1.00		2326	90.4	1919	91.2	1.00		0.25 (1.00)
AA	164	10.8	104	8.0	0.77	(0.59, 1.00)	248	9.6	185	8.8	0.94	(0.77, 1.15)	

	Pre/Peri-Menopause							Post-Menopause							
	Controls		Cases					Controls		Cases					
	N	(%)	N	(%)	OR [a]	(95% CI)		N	(%)	N	(%)	OR [a]	(95% CI)		p-int [b]
Wald-p [c]					0.21							1.00			
ERα (rs1801132)															
CC/CG	220	94.0	219	92.0	1.00			452	96.2	411	92.4	1.00			0.37 (1.00)
GG	14	6.0	19	8.0	1.36	(0.66,	2.78)	18	3.8	34	7.6	**2.14**	**(1.18,**	**3.87)**	
Wald-p [c]					0.65							**0.04**			
ERα (rs3798577)															
TT/TC	188	80.3	180	75.6	1.00			388	82.6	346	77.8	1.00			0.96 (1.00)
CC	46	19.7	58	24.4	1.32	(0.85,	2.06)	82	17.4	99	22.3	1.37	(0.98,	1.90)	
Wald-p [c]					0.57							0.22			
ERα (rs2046210)															
GG	107	45.7	92	38.7	1.00			215	45.7	189	42.5	1.00			0.50 (1.00)
GA/AA	127	54.3	146	61.3	1.36	(0.94,	1.97)	255	54.3	256	57.5	1.15	(0.88,	1.50)	
Wald-p [c]					0.51							0.78			
ERα (rs851984)															
GG	90	38.5	84	35.3	1.00			181	38.5	167	37.5	1.00			0.80 (1.00)
GA	110	47.0	118	49.6	1.15	(0.78,	1.71)	220	46.8	204	45.8	1.10	(0.75,	1.33)	
AA	34	14.5	36	15.1	1.12	(0.64,	1.96)	69	14.7	74	16.6	1.14	(0.77,	1.68)	
p-trend [c]					0.65							0.78			
ERα (rs6913578)															
AA	113	48.3	102	42.9	1.00			233	49.6	206	46.3	1.00			0.76 (1.00)
AC	103	44.0	111	46.6	1.21	(0.83,	1.78)	185	39.4	187	42.0	1.14	(0.86,	1.51)	
CC	18	7.7	25	10.5	1.57	(0.80,	3.05)	52	11.0	52	11.7	1.15	(0.75,	1.77)	
p-trend [c]					0.54							0.78			

[a] Odds Ratios adjusted for age, study, and genetic admixture

[b] Interaction p-value (SNP*menopause); Bonferroni p-value for adjustment for multiple comparisons shown in parentheses

[c] Wald (or trend) p-value within strata adjusted for multiple comparisons (MC) (Bonferroni-Holm step-down method), **bold** text indicates significance after MC adjustment

Table 13b. The association of TGF-β signaling genes and breast cancer stratified by proportion Native American ancestry *(Full table)*

Gene (SNP)	Genotype	Low (0-28%)				Moderate (29-70%)				High (71-100%)				
		Case/Control	OR (95% CI) [a]			Case/Control	OR (95% CI) [a]			Case/Control	OR (95% CI) [a]			p-int [b]
RUNX1 (rs7279383)														
	CC	1165/1229	1.00			1092/1285	1.00			390/584	1.00			**0.004 (0.04)**
	CG/GG	530/641	0.87	(0.76,	1.00)	288/419	**0.82**	**(0.69,**	**0.97)**	58/50	**1.75**	**(1.17,**	**2.63)**	
Wald-p [c]			0.41				**0.14**				**0.05**			
RUNX1 (rs2268288)														
	TT/TC	1613/1798	1.00			1356/1677	1.00			446/632	1.00			0.83 (1.00)
	CC	83/73	1.31	(0.95,	1.82)	24/27	1.09	(0.63,	1.91)	1/2	0.69	(0.06,	8.59)	
Wald-p [c]			0.69				1.00				1.00			
RUNX1 (rs2252585)														
	TT/TC	1565/1717	1.00			1137/1400	1.00			338/468	1.00			0.90 (1.00)
	CC	131/154	0.91	(0.71,	1.16)	242/304	0.98	(0.82,	1.19)	110/166	0.92	(0.69,	1.22)	
Wald-p [c]			1.00				1.00				1.00			
RUNX1 (rs11701453)														
	GG	1082/1210	1.00			978/1217	1.00			335/490	1.00			0.97 (1.00)
	GC	548/583	1.06	(0.92,	1.22)	374/451	1.05	(0.89,	1.23)	108/134	1.26	(0.94,	1.69)	
	CC	65/78	0.92	(0.65,	1.29)	27/36	0.94	(0.57,	1.56)	5/10	0.71	(0.24,	2.12)	
p-trend [c]			1.00				1.00				1.00			
RUNX1 (rs8127225)														
	TT	1229/1380	1.00			693/931	1.00			195/280	1.00			0.48 (1.00)
	TC/CC	467/490	1.05	(0.90,	1.22)	687/770	**1.19**	**(1.03,**	**1.38)**	253/352	1.01	(0.79,	1.29)	
Wald-p [c]			1.00				**0.12**				1.00			
RUNX1 (rs1474479)														
	GG/GA	1477/1591	1.00			1330/1651	1.00			444/632	1.00			0.17 (1.00)
	AA	218/280	0.87	(0.72,	1.05)	50/53	1.17	(0.79,	1.74)	4/2	3.35	(0.61,	18.57)	
Wald-p [c]			0.93				1.00				1.00			

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			p-int [b]
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		
RUNX1 (rs1883066)											
	GG	1277/1436	1.00		1168/1446	1.00		411/588	1.00		0.85 (1.00)
	GC/CC	419/434	1.08	(0.92, 1.26)	212/258	1.02	(0.84, 1.25)	37/46	1.20	(0.76, 1.89)	
Wald-p [c]			1.00			1.00			1.00		
RUNX1 (rs7279123)											
	CC	975/1024	1.00		919/1127	1.00		342/477	1.00		0.50 (1.00)
	CT	601/704	0.91	(0.79, 1.04)	412/508	0.99	(0.85, 1.16)	96/149	0.91	(0.68, 1.23)	
	TT	110/132	0.90	(0.69, 1.18)	43/66	0.80	(0.54, 1.19)	10/8	1.74	(0.67, 4.53)	
p-trend [c]			0.93			1.00			1.00		
RUNX2 (rs1321075)											
	CC	1178/1304	1.00		610/769	1.00		128/187	1.00		0.92 (1.00)
	CA	470/513	1.00	(0.86, 1.16)	612/764	1.01	(0.87, 1.17)	228/314	1.05	(0.79, 1.40)	
	AA	48/54	0.96	(0.65, 1.43)	157/171	1.14	(0.89, 1.45)	92/133	1.01	(0.71, 1.43)	
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs17209895)											
	TT	914/1018	1.00		1045/1230	1.00		395/556	1.00		0.26 (1.00)
	TC	659/731	1.01	(0.88, 1.16)	300/427	0.84	(0.71, 0.99)	48/76	0.89	(0.61, 1.32)	
	CC	122/122	1.11	(0.85, 1.45)	35/47	0.89	(0.57, 1.39)	5/2	4.07	(0.78, 21.40)	
p-trend [c]			1.00			0.57			1.00		
RUNX2 (rs2677108)											
	TT	568/637	1.00		328/383	1.00		77/109	1.00		0.94 (1.00)
	TC	833/906	1.02	(0.88, 1.18)	681/840	0.94	(0.79, 1.13)	213/302	0.99	(0.70, 1.40)	
	CC	295/328	0.99	(0.82, 1.21)	371/479	0.90	(0.74, 1.10)	158/223	0.99	(0.69, 1.42)	
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs2819854)											
	TT	406/444	1.00		367/477	1.00		125/192	1.00		0.85 (1.00)

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			p-int [b]
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		
	TC	865/948	1.01	(0.86, 1.19)	668/829	1.05	(0.88, 1.24)	219/303	1.13	(0.85, 1.51)	
	CC	425/478	0.99	(0.82, 1.19)	344/397	1.13	(0.92, 1.37)	104/139	1.13	(0.80, 1.60)	
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs2790093)											
	AA	747/833	1.00		671/808	1.00		191/276	1.00		0.78 (1.00)
	AG	763/851	1.01	(0.88, 1.16)	590/743	0.95	(0.82, 1.11)	199/288	1.06	(0.82, 1.38)	
	GG	186/187	1.12	(0.89, 1.40)	118/153	0.92	(0.71, 1.19)	58/70	1.24	(0.83, 1.85)	
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs9463090)											
	GG	1034/1166	1.00		935/1115	1.00		312/461	1.00		0.45 (1.00)
	GA	559/602	1.04	(0.90, 1.20)	392/528	0.89	(0.76, 1.04)	124/158	1.12	(0.85, 1.48)	
	AA	102/100	1.16	(0.87, 1.55)	52/60	1.04	(0.71, 1.53)	12/15	1.14	(0.52, 2.48)	
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs2396441)											
	CC	465/444	1.00		327/429	1.00		119/161	1.00		0.17 (1.00)
	CT	839/983	0.82	(0.70, 0.96)	699/880	1.04	(0.87, 1.24)	224/335	0.89	(0.67, 1.20)	
	TT	392/444	0.84	(0.69, 1.01)	354/394	1.18	(0.96, 1.45)	105/138	0.97	(0.68, 1.39)	
p-trend [c]			0.50			0.91			1.00		
RUNX2 (rs1316330)											
	GG	942/1050	1.00		970/1177	1.00		347/482	1.00		0.37 (1.00)
	GT	654/721	1.01	(0.88, 1.16)	385/484	0.96	(0.82, 1.13)	91/145	0.89	(0.66, 1.20)	
	TT	100/100	1.13	(0.84, 1.51)	25/41	0.74	(0.45, 1.22)	10/7	2.06	(0.77, 5.54)	
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs7750470)											
	TT	1044/1211	1.00		888/1073	1.00		320/433	1.00		0.14 (1.00)

Gene (SNP)	Genotype	Low (0-28%)				Moderate (29-70%)				High (71-100%)				p-int [b]
		Case/Control	OR (95% CI) [a]			Case/Control	OR (95% CI) [a]			Case/Control	OR (95% CI) [a]			
	TC	578/589	1.15	(1.00,	1.32)	437/554	0.96	(0.82,	1.12)	109/179	0.84	(0.63,	1.12)	
	CC	74/71	1.18	(0.84,	1.66)	55/77	0.88	(0.62,	1.26)	19/22	1.13	(0.60,	2.14)	
p-trend [c]			0.48				1.00				1.00			
RUNX2 (rs6930053)														
	CC	628/636	1.00			625/813	1.00			265/371	1.00			0.15 (1.00)
	CT	808/944	0.87	(0.75,	1.00)	596/731	1.06	(0.91,	1.23)	158/224	0.98	(0.75,	1.27)	
	TT	259/291	0.91	(0.74,	1.11)	159/160	1.29	(1.01,	1.65)	25/39	0.90	(0.53,	1.54)	
p-trend [c]			1.00				0.57				1.00			
RUNX2 (rs12208240)														
	GG	1442/1570	1.00			1029/1319	1.00			329/464	1.00			0.08 (0.80)
	GA	239/288	0.90	(0.74,	1.08)	328/366	1.15	(0.97,	1.37)	113/153	1.05	(0.79,	1.39)	
	AA	14/13	1.18	(0.55,	2.53)	22/19	1.53	(0.82,	2.85)	5/17	0.41	(0.15,	1.13)	
p-trend [c]			1.00				0.50				1.00			
RUNX2 (rs12209785)														
	AA/AG	1583/1774	1.00			1283/1586	1.00			397/571	1.00			0.47 (1.00)
	GG	111/97	1.29	(0.98,	1.72)	97/117	1.01	(0.76,	1.33)	51/62	1.19	(0.80,	1.77)	
Wald-p [c]			0.58				1.00				1.00			
RUNX2 (rs10948238)														
	CC/CT	1398/1603	1.00			1171/1440	1.00			381/559	1.00			0.13 (1.00)
	TT	295/267	**1.27**	**(1.06,**	**1.52)**	209/262	0.98	(0.81,	1.20)	67/75	1.34	(0.93,	1.91)	
Wald-p [c]			**0.12**				1.00				1.00			
RUNX2 (rs13201287)														
	GG/GA	1568/1767	1.00			1269/1562	1.00			389/570	1.00			0.10 (0.88)
	AA	128/104	1.39	(1.06,	1.82)	111/142	0.95	(0.73,	1.23)	59/64	1.36	(0.93,	1.99)	
Wald-p [c]			0.17				1.00				1.00			
RUNX2 (rs12333172)														

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		p-int [b]
	CC/CT	1616/1797	1.00		1341/1664	1.00		440/622	1.00		0.86 (1.00)
	TT	79/74	1.15 (0.83, 1.59)		39/40	1.23 (0.78, 1.92)		8/12	0.94 (0.38, 2.33)		
Wald-p [c]			1.00			1.00			1.00		
RUNX2 (rs1200428)											
	CC	991/1110	1.00		749/926	1.00		192/280	1.00		0.33 (1.00)
	CA	617/684	1.02 (0.89, 1.17)		546/653	1.03 (0.89, 1.20)		207/291	1.11 (0.85, 1.44)		
	AA	88/77	1.27 (0.93, 1.75)		85/125	0.82 (0.61, 1.10)		49/63	1.19 (0.78, 1.82)		
p-trend [c]			1.00			1.00			1.00		
RUNX2 (rs598953)											
	TT	680/722	1.00		478/585	1.00		108/173	1.00		0.18 (1.00)
	TA	791/909	0.93 (0.80, 1.07)		668/797	1.02 (0.87, 1.20)		224/323	1.13 (0.84, 1.53)		
	AA	225/240	0.99 (0.80, 1.23)		234/322	0.88 (0.72, 1.09)		116/138	1.40 (0.98, 1.98)		
p-trend [c]			1.00			1.00			0.92		
RUNX3 (rs2236850)											
	TT	515/602	1.00		476/613	1.00		198/271	1.00		0.57 (1.00)
	TC	844/890	1.10 (0.95, 1.28)		653/810	1.04 (0.89, 1.22)		189/283	0.93 (0.71, 1.20)		
	CC	335/374	1.04 (0.86, 1.26)		249/279	1.17 (0.95, 1.44)		61/80	0.99 (0.67, 1.47)		
p-trend [c]			0.94			0.86			1.00		
RUNX3 (rs9438876)											
	AA	373/430	1.00		445/567	1.00		198/268	1.00		0.34 (1.00)
	AG	846/888	1.09 (0.92, 1.29)		681/825	1.06 (0.90, 1.24)		186/291	0.89 (0.68, 1.15)		
	GG	476/553	0.99 (0.82, 1.19)		253/312	1.06 (0.86, 1.30)		64/75	1.11 (0.75, 1.64)		
p-trend [c]			0.94			1.00			1.00		
RUNX3 (rs7517302)											
	TT/TC	1359/1529	1.00		1170/1459	1.00		386/554	1.00		0.99 (1.00)
	CC	335/341	1.11 (0.94, 1.31)		209/243	1.09 (0.89, 1.33)		62/80	1.06 (0.73, 1.52)		

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			p-int [b]
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		
Wald-p [c]			0.94			1.00			1.00		
RUNX3 (rs906296)											
	CC	987/1113	1.00		873/1155	1.00		283/433	1.00		0.22 (1.00)
	CG/GG	707/758	1.05	(0.92, 1.20)	507/546	**1.23**	**(1.06, 1.43)**	165/201	1.24	(0.96, 1.61)	
Wald-p [c]			0.94			**0.04**			0.61		
RUNX3 (rs7551188)											
	CC	377/404	1.00		352/475	1.00		157/227	1.00		0.60 (1.00)
	CT	848/934	0.96	(0.81, 1.14)	699/821	1.16	(0.98, 1.38)	209/282	1.03	(0.79, 1.36)	
	TT	471/533	0.94	(0.78, 1.14)	327/403	1.10	(0.90, 1.35)	80/125	0.87	(0.61, 1.24)	
p-trend [c]			0.94			1.00			1.00		
RUNX3 (rs6688058)											
	GG/GA	1659/1842	1.00		1347/1666	1.00		440/619	1.00		0.37 (1.00)
	AA	37/29	1.50	(0.91, 2.45)	33/38	1.08	(0.68, 1.74)	8/15	0.74	(0.31, 1.77)	
Wald-p [c]			0.56			1.00			1.00		
RUNX3 (rs11249206)											
	TT	437/441	1.00		537/666	1.00		273/364	1.00		0.63 (1.00)
	TC	834/953	0.89	(0.75, 1.04)	633/781	1.01	(0.87, 1.19)	140/222	0.77	(0.59, 1.01)	
	CC	399/451	0.90	(0.74, 1.08)	191/227	1.06	(0.84, 1.32)	25/37	0.91	(0.53, 1.55)	
p-trend [c]			0.94			1.00			0.71		
RUNX3 (rs4478762)											
	GG/GA	1670/1856	1.00		1358/1681	1.00		442/625	1.00		0.48 (1.00)
	AA	26/15	2.01	(1.06, 3.81)	22/22	1.24	(0.68, 2.25)	6/8	1.09	(0.38, 3.18)	
Wald-p [c]			0.20			1.00			1.00		
TGF-β1 (rs1800469)											
	CC	772/866	1.00		379/516	1.00		102/141	1.00		0.42 (0.84)
	CT	692/759	1.01	(0.88, 1.17)	692/854	**1.11**	**(0.94, 1.31)**	206/299	0.95	(0.69, 1.30)	

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			p-int [b]
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		
	TT	190/216	0.96	(0.77, 1.20)	295/313	**1.29**	**(1.04, 1.58)**	133/193	0.92	(0.65, 1.30)	
p-trend [c]			1.00			**0.04**			0.94		
TGF-β1 (rs4803455)											
	CC	444/478	1.00		509/604	1.00		240/318	1.00		0.63 (0.84)
	CA/AA	1173/1339	0.95	(0.82, 1.11)	670/932	**0.86**	**(0.73, 1.00)**	189/290	0.91	(0.71, 1.17)	
Wald-p [c]			1.00			**0.05**			0.94		
TGF-βRI (rs6478974)											
	TT/TA	1298/1467	1.00		1183/1485	1.00		403/579	1.00		0.95 (1.00)
	AA	397/404	1.12	(0.96, 1.31)	197/218	1.15	(0.93, 1.41)	45/55	1.25	(0.82, 1.90)	
Wald-p [c]			0.65			0.20			1.00		
TGF-βRI (rs1571590)											
	AA	1080/1212	1.00		1096/1328	1.00		404/571	1.00		0.24 (1.00)
	AG	561/589	1.07	(0.93, 1.24)	267/362	0.89	(0.74, 1.06)	44/63	0.98	(0.65, 1.48)	
	GG	55/69	0.91	(0.63, 1.31)	17/14	1.45	(0.71, 2.96)	0/0	-	- -	
p-trend [c]			1.00			0.74			1.00		
TGF-βRI (rs1013186)											
	GG	1080/1210	1.00		1095/1329	1.00		403/569	1.00		0.50 (1.00)
	GA	561/592	1.06	(0.92, 1.23)	268/360	0.90	(0.75, 1.07)	45/64	0.98	(0.65, 1.47)	
	AA	55/69	0.91	(0.63, 1.31)	17/15	1.35	(0.67, 2.72)	0/1	-	- -	
p-trend [c]			1.00			0.74			1.00		
TGF-βRI (rs11568785)											
	AA/AG	1682/1856	1.00		1373/1703	1.00		448/633	1.00		0.19 (0.44)
	GG	14/15	1.09	(0.52, 2.27)	7/1	8.28	(1.02, 67.40)	0/1	-	- -	
Wald-p [c]			1.00			0.23			1.00		
TGF-βRI (rs10733710)											
	GG/GA	1620/1777	1.00		1232/1495	1.00		369/513	1.00		0.95 (1.00)

Gene (SNP)	Genotype	Low (0-28%)			Moderate (29-70%)			High (71-100%)			
		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		Case/Control	OR (95% CI) [a]		p-int [b]
	AA	75/93	0.88	(0.64, 1.20)	148/209	0.87	(0.69, 1.09)	79/121	0.89	(0.65, 1.22)	
Wald-p [c]			1.00			0.74			1.00		

[a] Odds Ratios adjusted for age and study

[b] Interaction p-value (SNP*admixture); Bonferroni-Holm p-value for multiple comparisons shown in parenthesis; **bold** text indicates significance after MC adjustment

[c] Wald (or trend) p-value within strata adjusted for multiple comparisons (MC) by admixture strata (Bonferroni-Holm step-down method), **bold** text indicates significance ($p \leq 0.05$) or suggestive of an association ($p \leq 0.15$) after MC adjustment

Table 15b. The Association of TGF-β signaling genes and breast cancer defined by ER/PR status* *(Full table)*

		Controls	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
Gene (SNP)	Genotype	**N**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**p [b]**
RUNX1 (rs7279383)															
	CC	2249	927	1.00		167	1.00		32	1.00		306	1.00		0.75
	CG/GG	965	351	0.87	(0.75, 1.00)	66	0.92	(0.69, 1.24)	13	0.99	(0.51, 1.91)	100	0.77	(0.60, 0.98)	
	Wald-p [c]			0.42			1.00			1.00			0.26		
RUNX1 (rs2268288)															
	TT/TC	3122	1234	1.00		225	1.00		42	1.00		391	1.00		1.00
	CC	93	45	1.19	(0.83, 1.71)	8	1.18	(0.57, 2.47)	3	2.59	(0.78, 8.62)	15	1.32	(0.76, 2.32)	
	Wald-p [c]			1.00			1.00			0.98			1.00		
RUNX1 (rs2252585)															
	TT/TC	2822	1142	1.00		200	1.00		39	1.00		358	1.00		1.00
	CC	393	137	0.89	(0.72, 1.09)	33	1.19	(0.81, 1.76)	6	0.99	(0.41, 2.38)	47	0.90	(0.65, 1.25)	
	Wald-p [c]			1.00			1.00			1.00			1.00		
RUNX1 (rs11701453)															
	GG	2179	854	1.00		168	1.00		31	1.00		266	1.00		1.00
	GC	931	382	1.03	(0.89, 1.19)	60	0.83	(0.61, 1.13)	14	1.09	(0.58, 2.07)	130	1.16	(0.92, 1.45)	
	CC	105	42	0.99	(0.68, 1.43)	5	0.62	(0.25, 1.53)	0	0.00	(0.00, 0.00)	9	0.71	(0.35, 1.42)	
	p-trend [c]			1.00			0.77			1.00			1.00		
RUNX1 (rs8127225)															
	TT	2132	821	1.00		137	1.00		28	1.00		262	1.00		0.67
	TC/CC	1078	458	1.14	(0.99, 1.31)	96	1.40	(1.06, 1.85)	17	1.08	(0.58, 2.02)	144	1.04	(0.83, 1.30)	
	Wald-p [c]			0.48			0.12			1.00			1.00		
RUNX1 (rs1474479)															
	GG/GA	2896	1150	1.00		216	1.00		40	1.00		375	1.00		1.00
	AA	319	129	0.97	(0.78, 1.20)	17	0.69	(0.41, 1.15)	5	1.35	(0.52, 3.54)	30	0.77	(0.52, 1.14)	
	Wald-p [c]			1.00			0.77			1.00			1.00		

Gene (SNP)	Genotype	Controls N	ER+/PR+ N	ER+/PR+ OR [a]	ER+/PR+ (95% CI)	ER+/PR- N	ER+/PR- OR [a]	ER+/PR- (95% CI)	ER-/PR+ N	ER-/PR+ OR [a]	ER-/PR+ (95% CI)	ER-/PR- N	ER-/PR- OR [a]	ER-/PR- (95% CI)	p [b]
RUNX1 (rs1883066)															
	GG	2569	1001	1.00		189	1.00		38	1.00		325	1.00		1.00
	GC/CC	645	278	1.09	(0.93, 1.28)	44	0.92	(0.66, 1.30)	7	0.79	(0.35, 1.79)	81	1.03	(0.79, 1.33)	
	Wald-p [c]			1.00			1.00			1.00			1.00		
RUNX1 (rs7279123)															
	CC	1935	730	1.00		160	1.00		31	1.00		249	1.00		0.06
	CT	1083	466	1.12	(0.97, 1.29)	65	**0.72**	**(0.53, 0.97)**	14	0.83	(0.44, 1.57)	135	0.98	(0.78, 1.22)	
	TT	184	79	1.10	(0.83, 1.46)	7	**0.44**	**(0.21, 0.96)**	0	0.00	(0.00, 0.00)	20	0.88	(0.55, 1.43)	
	p-trend [c]			0.83			**0.03**			1.00			1.00		
RUNX2 (rs1321075)															
	CC	1918	741	1.00		139	1.00		23	1.00		240	1.00		1.00
	CA	1102	456	1.11	(0.97, 1.28)	80	1.00	(0.75, 1.34)	19	1.33	(0.71, 2.51)	127	0.90	(0.72, 1.14)	
	AA	195	81	1.17	(0.89, 1.55)	14	1.00	(0.56, 1.80)	3	1.18	(0.34, 4.09)	39	1.56	(1.06, 2.29)	
	p-trend [c]			0.97			1.00			1.00			0.34		
RUNX2 (rs17209895)															
	TT	2002	816	1.00		154	1.00		23	1.00		244	1.00		0.94
	TC	1054	408	0.92	(0.80, 1.06)	67	0.82	(0.60, 1.11)	21	1.92	(1.04, 3.54)	139	1.12	(0.89, 1.41)	
	CC	159	55	0.82	(0.60, 1.13)	12	1.00	(0.54, 1.85)	1	0.65	(0.09, 4.89)	23	1.26	(0.79, 2.00)	
	p-trend [c]			1.00			1.00			1.00			0.61		
RUNX2 (rs2677108)															
	TT	952	378	1.00		74	1.00		7	1.00		119	1.00		1.00
	TC	1559	646	1.06	(0.91, 1.24)	115	0.95	(0.70, 1.29)	29	2.48	(1.08, 5.71)	186	0.96	(0.75, 1.23)	
	CC	702	255	0.95	(0.79, 1.15)	44	0.82	(0.55, 1.21)	9	1.64	(0.60, 4.46)	101	1.14	(0.85, 1.52)	
	p-trend [c]			1.00			1.00			1.00			0.61		
RUNX2 (rs2819854)															
	TT	819	316	1.00		51	1.00		12	1.00		115	1.00		1.00

		Controls	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
Gene (SNP)	Genotype	**N**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**p [b]**
	TC	1599	643	1.03	(0.88, 1.21)	120	1.19	(0.85, 1.67)	27	1.16	(0.58, 2.31)	188	0.84	(0.65, 1.07)	
	CC	795	319	1.02	(0.85, 1.22)	62	1.22	(0.83, 1.79)	6	0.51	(0.19, 1.38)	103	0.91	(0.68, 1.21)	
	p-trend [c]			1.00			1.00			1.00			0.61		
RUNX2 (rs2790093)															
	AA	1450	579	1.00		105	1.00		27	1.00		187	1.00		1.00
	AG	1458	574	0.98	(0.86, 1.12)	109	1.02	(0.77, 1.35)	13	0.49	(0.25, 0.96)	180	0.97	(0.78, 1.21)	
	GG	307	126	1.02	(0.81, 1.28)	19	0.84	(0.51, 1.39)	5	0.96	(0.37, 2.52)	39	1.02	(0.71, 1.48)	
	p-trend [c]			1.00			1.00			1.00			0.96		
RUNX2 (rs9463090)															
	GG	2051	818	1.00		155	1.00		25	1.00		238	1.00		0.06
	GA	1016	410	1.02	(0.88, 1.17)	70	0.93	(0.70, 1.25)	17	1.34	(0.72, 2.50)	130	**1.10**	**(0.87, 1.38)**	
	AA	144	49	0.83	(0.59, 1.16)	8	0.72	(0.35, 1.50)	3	1.81	(0.54, 6.10)	38	**2.31**	**(1.57, 3.39)**	
	p-trend [c]			1.00			1.00			1.00			**0.009**		
RUNX2 (rs2396441)															
	CC	785	318	1.00		64	1.00		11	1.00		110	1.00		1.00
	CT	1684	625	0.92	(0.78, 1.08)	113	0.82	(0.60, 1.13)	16	0.68	(0.31, 1.46)	207	0.88	(0.69, 1.13)	
	TT	745	336	1.11	(0.93, 1.34)	56	0.92	(0.63, 1.34)	18	1.70	(0.80, 3.62)	89	0.85	(0.63, 1.14)	
	p-trend [c]			1.00			1.00			1.00			0.61		
RUNX2 (rs1316330)															
	GG	1971	773	1.00		145	1.00		29	1.00		261	1.00		1.00
	GT	1109	459	1.03	(0.90, 1.19)	78	0.95	(0.71, 1.26)	14	0.94	(0.49, 1.79)	128	0.90	(0.72, 1.13)	
	TT	133	47	0.87	(0.62, 1.23)	10	1.02	(0.52, 2.00)	2	1.17	(0.27, 5.01)	17	1.02	(0.60, 1.72)	
	p-trend [c]			1.00			1.00			1.00			0.61		
RUNX2 (rs7750470)															
	TT	2060	830	1.00		145	1.00		31	1.00		242	1.00		1.00
	TC	1024	405	0.98	(0.85, 1.13)	78	1.10	(0.82, 1.46)	13	0.82	(0.43, 1.58)	142	1.17	(0.94, 1.46)	

		Controls	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
Gene (SNP)	Genotype	**N**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**p [b]**
	CC	131	44	0.83	(0.59, 1.18)	10	1.10	(0.56, 2.13)	1	0.47	(0.06, 3.51)	22	1.39	(0.86, 2.23)	
	p-trend [c]			1.00			1.00			1.00			0.34		
RUNX2 (rs6930053)															
	CC	1280	493	1.00		86	1.00		18	1.00		183	1.00		0.96
	CT	1516	598	1.01	(0.88, 1.16)	118	1.16	(0.87, 1.54)	20	0.99	(0.52, 1.89)	181	0.85	(0.68, 1.06)	
	TT	419	187	1.13	(0.92, 1.38)	29	1.03	(0.66, 1.59)	7	1.28	(0.53, 3.11)	42	0.71	(0.50, 1.02)	
	p-trend [c]			1.00			1.00			1.00			0.27		
RUNX2 (rs12208240)															
	GG	2599	1020	1.00		189	1.00		39	1.00		325	1.00		1.00
	GA	583	241	1.07	(0.91, 1.27)	42	1.00	(0.70, 1.41)	6	0.64	(0.27, 1.52)	74	0.99	(0.75, 1.30)	
	AA	33	17	1.33	(0.73, 2.40)	2	0.79	(0.19, 3.33)	0	0.00	(0.00, 0.00)	5	1.18	(0.45, 3.06)	
	p-trend [c]			1.00			1.00			1.00			0.94		
RUNX2 (rs12209785)															
	AA/AG	3024	1196	1.00		220	1.00		42	1.00		372	1.00		1.00
	GG	190	82	1.10	(0.84, 1.44)	13	0.92	(0.52, 1.65)	3	1.20	(0.37, 3.93)	34	1.49	(1.01, 2.18)	
	Wald-p [c]			1.00			1.00			1.00			0.27		
RUNX2 (rs10948238)															
	CC/CT	2742	1072	1.00		199	1.00		36	1.00		328	1.00		0.94
	TT	470	205	1.11	(0.93, 1.33)	33	0.97	(0.66, 1.41)	9	1.47	(0.70, 3.09)	78	1.39	(1.07, 1.82)	
	Wald-p [c]			1.00			1.00			1.00			0.12		
RUNX2 (rs13201287)															
	GG/GA	3003	1185	1.00		221	1.00		41	1.00		370	1.00		1.00
	AA	212	94	1.13	(0.88, 1.46)	12	0.75	(0.41, 1.37)	4	1.43	(0.51, 4.05)	36	1.40	(0.96, 2.03)	
	Wald-p [c]			1.00			1.00			1.00			0.34		
RUNX2 (rs12333172)															
	CC/CT	3109	1238	1.00		222	1.00		43	1.00		378	1.00		0.06

		Controls	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
Gene (SNP)	Genotype	**N**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**p [b]**
	TT	106	40	0.92	(0.63, 1.33)	11	1.44	(0.76, 2.73)	2	1.36	(0.32, 5.72)	28	**2.12**	**(1.37, 3.27)**	
	Wald-p [c]			1.00			1.00			1.00			**0.007**		
RUNX2 (rs1200428)															
	CC	1818	720	1.00		130	1.00		29	1.00		233	1.00		1.00
	CA	1220	491	1.02	(0.89, 1.17)	96	1.10	(0.83, 1.44)	14	0.71	(0.37, 1.36)	146	0.93	(0.75, 1.16)	
	AA	177	68	0.98	(0.73, 1.31)	7	0.53	(0.25, 1.16)	2	0.73	(0.17, 3.10)	27	1.20	(0.78, 1.84)	
	p-trend [c]			1.00			1.00			1.00			0.93		
RUNX2 (rs598953)															
	TT	1175	485	1.00		92	1.00		15	1.00		147	1.00		1.00
	TA	1539	607	0.96	(0.83, 1.11)	112	0.92	(0.69, 1.23)	24	1.23	(0.64, 2.36)	204	1.07	(0.85, 1.34)	
	AA	501	187	0.92	(0.75, 1.12)	29	0.74	(0.48, 1.14)	6	0.91	(0.35, 2.37)	55	0.87	(0.63, 1.21)	
	p-trend [c]			1.00			1.00			1.00			0.63		
RUNX3 (rs2236850)															
	TT	1073	401	1.00		77	1.00		11	1.00		105	1.00		0.17
	TC	1526	620	1.08	(0.93, 1.26)	112	1.02	(0.76, 1.38)	15	**0.96**	**(0.44, 2.10)**	216	1.44	(1.13, 1.85)	
	CC	609	257	1.12	(0.93, 1.35)	44	1.00	(0.68, 1.46)	18	**2.88**	**(1.35, 6.16)**	83	1.39	(1.02, 1.88)	
	p-trend [c]			0.64			1.00			**0.03**			0.09		
RUNX3 (rs9438876)															
	AA	871	325	1.00		63	1.00		8	1.00		100	1.00		0.40
	AG	1527	634	1.10	(0.94, 1.28)	120	1.08	(0.79, 1.49)	19	1.37	(0.60, 3.15)	192	1.09	(0.84, 1.41)	
	GG	817	319	1.01	(0.85, 1.22)	50	0.83	(0.57, 1.23)	18	2.52	(1.08, 5.87)	113	1.21	(0.91, 1.62)	
	p-trend [c]			0.86			1.00			0.09			0.75		
RUNX3 (rs7517302)															
	TT/TC	2664	1035	1.00		191	1.00		32	1.00		324	1.00		**0.03**
	CC	548	242	1.13	(0.96, 1.34)	41	1.05	(0.74, 1.50)	13	2.02	(1.05, 3.88)	82	**1.25**	**(0.96, 1.62)**	
	Wald-p [c]			0.64			1.00			0.08			**0.03**		

Gene (SNP)	Genotype	Controls N	ER+/PR+ N	ER+/PR+ OR [a]	ER+/PR+ (95% CI)	ER+/PR- N	ER+/PR- OR [a]	ER+/PR- (95% CI)	ER-/PR+ N	ER-/PR+ OR [a]	ER-/PR+ (95% CI)	ER-/PR- N	ER-/PR- OR [a]	ER-/PR- (95% CI)	p [b]
RUNX3 (rs906296)															
	CC	2009	748	1.00		134	1.00		27	1.00		247	1.00		0.40
	CG/GG	1203	530	1.18	(1.03, 1.34)	98	1.22	(0.93, 1.61)	18	1.16	(0.63, 2.12)	159	1.09	(0.88, 1.35)	
	Wald-p [c]			0.09			0.86			0.63			1.00		
RUNX3 (rs7551188)															
	CC	781	296	1.00		50	1.00		8	1.00		104	1.00		0.94
	CT	1569	650	1.08	(0.92, 1.27)	129	1.27	(0.91, 1.79)	24	1.48	(0.66, 3.31)	198	0.95	(0.73, 1.22)	
	TT	860	333	1.01	(0.84, 1.21)	53	0.97	(0.65, 1.44)	11	1.25	(0.50, 3.15)	103	0.90	(0.67, 1.21)	
	p-trend [c]			0.97			1.00			1.00			1.00		
RUNX3 (rs6688058)															
	GG/GA	3155	1247	1.00		231	1.00		45	1.00		397	1.00		0.46
	AA	60	32	1.37	(0.88, 2.11)	2	0.46	(0.11, 1.88)	0	0.00	(0.00, 0.00)	9	1.20	(0.59, 2.45)	
	Wald-p [c]			0.64			1.00			1.00			1.00		
RUNX3 (rs11249206)															
	TT	951	394	1.00		63	1.00		12	1.00		134	1.00		0.47
	TC	1560	611	0.92	(0.79, 1.07)	120	1.15	(0.84, 1.59)	22	1.16	(0.57, 2.38)	185	0.85	(0.67, 1.08)	
	CC	649	248	0.88	(0.73, 1.07)	48	1.12	(0.75, 1.66)	11	1.46	(0.63, 3.38)	84	0.94	(0.70, 1.27)	
	p-trend [c]			0.64			1.00			1.00			1.00		
RUNX3 (rs4478762)															
	GG/GA	3178	1253	1.00		232	1.00		45	1.00		400	1.00		0.40
	AA	35	26	1.90	(1.14, 3.18)	1	0.39	(0.05, 2.88)	0	0.00	(0.00, 0.00)	6	1.35	(0.56, 3.23)	
	P-trend			0.09			1.00			1.00			1.00		
TGF-β1 (rs1800469)															
	CC	1269	511	1.00		81	1.00		14	1.00		145	1.00		0.56
	CT	1425	544	0.96	(0.83, 1.11)	106	1.19	(0.88, 1.60)	21	1.22	(0.61, 2.43)	191	1.13	(0.90, 1.43)	
	TT	469	191	1.05	(0.86, 1.28)	42	1.43	(0.96, 2.11)	9	1.60	(0.68, 3.80)	64	1.16	(0.84, 1.59)	

Gene (SNP)	Genotype	Controls N	ER+/PR+ N	ER+/PR+ OR [a]	ER+/PR+ (95% CI)	ER+/PR- N	ER+/PR- OR [a]	ER+/PR- (95% CI)	ER-/PR+ N	ER-/PR+ OR [a]	ER-/PR+ (95% CI)	ER-/PR- N	ER-/PR- OR [a]	ER-/PR- (95% CI)	p [b]
	p-trend [c]			0.85			0.15			0.58			0.57		
TGF-β1 (rs4803455)															
	CC	927	374	1.00		64	1.00		13	1.00		113	1.00		0.61
	CA/AA	2041	762	0.90	(0.77, 1.04)	134	0.92	(0.67, 1.25)	26	1.00	(0.50, 1.96)	223	0.89	(0.70, 1.14)	
	Wald-p [c]			0.29			0.58			0.99			0.57		
TGF-βR1 (rs6478974)															
	TT/TA	2634	1034	1.00		190	1.00		37	1.00		323	1.00		1.00
	AA	580	244	1.05	(0.89, 1.25)	43	1.03	(0.73, 1.45)	8	1.08	(0.50, 2.35)	83	1.21	(0.94, 1.58)	
	Wald-p [c]			1.00			1.00			0.85			0.43		
TGF-βR1 (rs1571590)															
	AA	2277	892	1.00		161	1.00		36	1.00		288	1.00		1.00
	AG	859	360	1.04	(0.90, 1.21)	67	1.10	(0.82, 1.48)	9	0.70	(0.33, 1.47)	106	1.00	(0.79, 1.27)	
	GG	78	27	0.83	(0.53, 1.30)	5	0.86	(0.34, 2.18)	0	0.00	(0.00, 0.00)	12	1.30	(0.69, 2.43)	
	p-trend [c]			1.00			1.00			0.57			1.00		
TGF-βR1 (rs1013186)															
	GG	2275	892	1.00		160	1.00		36	1.00		288	1.00		1.00
	GA	862	360	1.04	(0.90, 1.20)	68	1.12	(0.83, 1.51)	9	0.70	(0.33, 1.46)	106	0.99	(0.78, 1.26)	
	AA	78	27	0.83	(0.53, 1.30)	5	0.87	(0.35, 2.20)	0	0.00	(0.00, 0.00)	12	1.30	(0.69, 2.43)	
	p-trend [c]			1.00			1.00			0.57			1.00		
TGF-βR1 (rs11568785)															
	AA/AG	3199	1272	1.00		229	1.00		44	1.00		405	1.00		0.27
	GG	16	7	1.08	(0.44, 2.63)	4	3.55	(1.17, 10.80)	1	5.70	(0.73, 44.6)	1	0.55	(0.07, 4.21)	
	Wald-p [c]			1.00			0.10			0.39			1.00		
TGF-βR1 (rs10733710)															
	GG/GA	2938	1195	1.00		222	1.00		41	1.00		380	1.00		0.21
	AA	276	84	0.78	(0.60, 1.00)	11	0.53	(0.28, 0.99)	4	0.92	(0.32, 2.61)	25	0.67	(0.44, 1.03)	

		Controls	ER+/PR+			ER+/PR-			ER-/PR+			ER-/PR-			
Gene (SNP)	Genotype	**N**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**N**	**OR [a]**	**(95% CI)**	**p [b]**
	Wald-p [c]			0.21			0.14			0.87			0.27		

* ER/PR data were compared with 3,214 controls from sites where cases have ER/PR data Mexico data is excluded because they do not have data for ER/PR status (n=1,810)

[a] Odds ratios (OR) and 95% confidence intervals (CI) adjusted for age, study, and genetic admixture.

[b] Wald-p for overall model, Bonferroni-Holm p-value adjustment for MC shown

[c] Wald-p (dominant/recessive) or p-trend (additive) adjusted for multiple comparisons within strata; **bold text** indicates significance remained after MC adjustment

Table 17b. TGF-β signaling and ERα SNPs and breast cancer defined by ER status, stratified by menopausal status (*Full table*)

	Pre-menopausal								Post-menopausal								
	Controls [a]	ER+				ER-				ER+				ER-			
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
RUNX1 (rs7279383)																		
CC	2166	351	1.00			145	1.00			694	1.00			175	1.00			1.00
CG	864	130	0.88	(0.69,	1.13)	51	0.86	(0.61,	1.23)	242	0.87	(0.72,	1.03)	45	**0.65**	**(0.46,**	**0.92)**	
GG	71	11	0.92	(0.44,	1.90)	3	0.64	(0.19,	2.15)	23	1.01	(0.61,	1.68)	4	**0.76**	**(0.27,**	**2.13)**	
p-trend			1.00				1.00				1.00				**0.16**			
RUNX1 (rs2268288)																		
TT/TC	3010	477	1.00			194	1.00			925	1.00			211	1.00			0.59
CC	91	15	0.83	(0.45,	1.53)	5	0.70	(0.27,	1.80)	34	1.37	(0.88,	2.13)	13	**2.47**	**(1.31,**	**4.65)**	
Wald-p			1.00				1.00				1.00				**0.04**			
RUNX1 (rs2252585)																		
TT/TC	2719	434	1.00			172	1.00			855	1.00			201	1.00			1.00
CC	382	58	0.92	(0.66,	1.28)	27	1.01	(0.65,	1.59)	104	0.92	(0.72,	1.18)	22	0.79	(0.50,	1.26)	
Wald-p			1.00				1.00				1.00				1.00			
RUNX1 (rs11701453)																		
GG	2101	341	1.00			130	1.00			645	1.00			148	1.00			1.00
GC	900	134	0.97	(0.76,	1.24)	64	1.25	(0.90,	1.73)	286	0.99	(0.84,	1.18)	71	1.07	(0.79,	1.45)	
CC	100	17	1.12	(0.61,	2.06)	4	0.71	(0.25,	2.04)	27	0.81	(0.51,	1.29)	5	0.68	(0.27,	1.72)	
p-trend			1.00				1.00				1.00				1.00			
RUNX1 (rs8127225)																		
TT	2062	303	1.00			121	1.00			615	1.00			153	1.00			1.00
TC	914	167	1.32	(1.04,	1.68)	69	1.29	(0.92,	1.79)	312	1.15	(0.97,	1.36)	64	0.88	(0.65,	1.21)	
CC	120	22	1.20	(0.71,	2.04)	9	1.11	(0.52,	2.33)	32	0.98	(0.64,	1.52)	7	0.74	(0.33,	1.66)	
p-trend			0.34				1.00				1.00				1.00			
RUNX1 (rs1474479)																		

	Pre-menopausal								Post-menopausal							
	Controls [a]	ER+			ER-				ER+				ER-			
	N	N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
GG/GA	2788	446	1.00		181	1.00			866	1.00			208	1.00			1.00
AA	313	46	1.08	(0.74, 1.58)	17	1.02	(0.59, 1.77)		93	0.84	(0.65, 1.09)		16	0.68	(0.40, 1.17)		
Wald-p			1.00			1.00				1.00				0.99			
RUNX1 (rs1883066)																	
GG	2470	390	1.00		158	1.00			750	1.00			180	1.00			1.00
GC/CC	630	102	1.10	(0.84, 1.44)	41	1.13	(0.77, 1.65)		209	1.03	(0.85, 1.24)		44	0.93	(0.66, 1.32)		
Wald-p			1.00			1.00				1.00				1.00			
RUNX1 (rs7279123)																	
CC	1862	290	1.00		119	1.00			564	1.00			140	1.00			1.00
CT	1049	170	1.03	(0.81, 1.29)	70	1.04	(0.75, 1.43)		337	1.05	(0.89, 1.24)		72	0.93	(0.69, 1.26)		
TT	178	31	1.36	(0.85, 2.18)	9	0.99	(0.47, 2.07)		54	0.90	(0.64, 1.25)		11	0.79	(0.41, 1.50)		
p-trend			1.00			1.00				1.00				1.00			
RUNX2 (rs1321075)																	
CC	1855	285	1.00		119	1.00			564	1.00			128	1.00			1.00
CA	1059	178	1.03	(0.81, 1.30)	63	0.82	(0.58, 1.15)		338	1.14	(0.96, 1.35)		75	1.08	(0.79, 1.47)		
AA	187	28	0.95	(0.59, 1.54)	17	1.29	(0.71, 2.32)		57	1.15	(0.82, 1.61)		21	1.62	(0.97, 2.71)		
p-trend			1.00			1.00				1.00				1.00			
RUNX2 (rs17209895)																	
TT	1926	316	1.00		113	1.00			612	1.00			134	1.00			1.00
TC	1017	160	1.00	(0.79, 1.27)	75	1.42	(1.02, 1.98)		298	0.86	(0.72, 1.02)		80	1.12	(0.83, 1.51)		
CC	158	16	0.61	(0.34, 1.09)	11	1.39	(0.69, 2.77)		49	0.93	(0.65, 1.33)		10	0.89	(0.45, 1.75)		
p-trend			1.00			0.37				1.00				1.00			
RUNX2 (rs2677108)																	
TT	920	146	1.00		55	1.00			289	1.00			62	1.00			1.00
TC	1503	257	1.14	(0.89, 1.46)	98	1.15	(0.80, 1.65)		478	1.00	(0.84, 1.20)		106	1.02	(0.73, 1.42)		

	Pre-menopausal								Post-menopausal								
	Controls [a]	**ER+**				**ER-**				**ER+**				**ER-**			
	N	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**p-int[c]**
CC	676	89	0.83	(0.60,	1.14)	46	1.14	(0.74,	1.75)	192	0.95	(0.76,	1.18)	56	1.19	(0.81,	1.75)	
p-trend			1.00				1.00				1.00				1.00			
RUNX2 (rs2819854)																		
TT	787	114	1.00			52	1.00			234	1.00			67	1.00			1.00
TC	1543	253	1.11	(0.85,	1.46)	95	0.88	(0.61,	1.28)	482	1.05	(0.87,	1.27)	106	0.85	(0.62,	1.18)	
CC	769	125	1.06	(0.78,	1.44)	52	0.92	(0.61,	1.41)	242	1.07	(0.86,	1.33)	51	0.83	(0.56,	1.22)	
p-trend			1.00				1.00				1.00				1.00			
RUNX2 (rs2790093)																		
AA	1403	230	1.00			102	1.00			429	1.00			102	1.00			1.00
AG	1402	224	0.95	(0.76,	1.19)	79	0.77	(0.56,	1.06)	435	1.02	(0.87,	1.20)	101	1.01	(0.76,	1.36)	
GG	296	38	0.77	(0.51,	1.15)	18	0.86	(0.50,	1.49)	95	1.04	(0.79,	1.36)	21	0.99	(0.60,	1.63)	
p-trend			1.00				1.00				1.00				1.00			
RUNX2 (rs9463090)																		
GG	1976	319	1.00			115	1.00			614	1.00			131	1.00			1.00
GA	978	152	0.89	(0.71,	1.12)	66	1.05	(0.76,	1.47)	309	1.08	(0.91,	1.28)	73	**1.19**	**(0.88,**	**1.62)**	
AA	143	21	1.04	(0.60,	1.79)	18	2.51	(1.39,	4.54)	35	0.71	(0.48,	1.06)	20	**1.97**	**(1.17,**	**3.30)**	
p-trend			1.00				0.34				1.00				**0.13**			
RUNX2 (rs2396441)																		
CC	755	127	1.00			51	1.00			233	1.00			57	1.00			1.00
CT	1620	230	0.82	(0.63,	1.07)	95	0.84	(0.58,	1.22)	487	0.99	(0.82,	1.20)	116	0.97	(0.69,	1.36)	
TT	725	135	1.13	(0.84,	1.53)	53	1.09	(0.71,	1.66)	239	1.06	(0.85,	1.32)	51	0.92	(0.62,	1.38)	
p-trend			1.00				1.00				1.00				1.00			
RUNX2 (rs1316330)																		
GG	1903	309	1.00			126	1.00			569	1.00			146	1.00			1.00
GT	1069	170	0.98	(0.78,	1.23)	66	0.99	(0.71,	1.37)	351	1.07	(0.90,	1.26)	68	0.83	(0.61,	1.13)	

	Pre-menopausal								Post-menopausal								
	Controls [a]	ER+				ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
TT	127	13	0.57	(0.30,	1.08)	7	0.82	(0.36, 1.86)	39	1.02	(0.69,	1.53)	10	1.09	(0.55,	2.17)	
p-trend			1.00				1.00			1.00				1.00			
RUNX2 (rs7750470)																	
TT	1991	302	1.00			122	1.00		633	1.00			136	1.00			1.00
TC	989	169	1.16	(0.92,	1.46)	69	1.16	(0.84, 1.60)	296	0.94	(0.79,	1.11)	75	1.11	(0.82,	1.49)	
CC	121	21	1.08	(0.63,	1.85)	8	0.99	(0.45, 2.15)	30	0.81	(0.52,	1.25)	13	1.61	(0.87,	2.98)	
p-trend			1.00				1.00			1.00				1.00			
RUNX2 (rs6930053)																	
CC	1232	185	1.00			88	1.00		363	1.00			97	1.00			1.00
CT	1462	230	1.05	(0.83,	1.33)	87	0.86	(0.62, 1.19)	467	1.06	(0.89,	1.25)	103	0.90	(0.67,	1.20)	
TT	407	76	1.38	(0.99,	1.94)	24	0.96	(0.58, 1.58)	129	0.98	(0.77,	1.26)	24	0.71	(0.44,	1.14)	
p-trend			0.79				1.00			1.00				1.00			
RUNX2 (rs12208240)																	
GG	2514	395	1.00			158	1.00		765	1.00			182	1.00			1.00
GA/AA	587	97	1.02	(0.78,	1.34)	41	1.05	(0.72, 1.53)	193	1.13	(0.93,	1.38)	40	0.95	(0.66,	1.36)	
Wald-p			1.00				1.00			1.00				1.00			
RUNX2 (rs12209785)																	
AA/AG	2917	468	1.00			183	1.00		896	1.00			208	1.00			1.00
GG	183	24	0.82	(0.50,	1.32)	16	1.43	(0.80, 2.53)	62	1.10	(0.80,	1.51)	16	1.25	(0.73,	2.16)	
Wald-p							1.00			1.00				1.00			
RUNX2 (rs10948238)																	
CC/CT	2643	416	1.00			159	1.00		809	1.00			185	1.00			1.00
TT	455	76	1.10	(0.81,	1.48)	40	1.52	(1.03, 2.24)	148	1.04	(0.83,	1.28)	39	1.21	(0.84,	1.75)	
Wald-p			1.00				0.34							1.00			
RUNX2 (rs13201287)																	

	Pre-menopausal								Post-menopausal								
	Controls [a]	**ER+**				**ER-**				**ER+**				**ER-**			
	N	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**p-int[c]**
GG/GA	2896	464	1.00			180	1.00			888	1.00			208	1.00			1.00
AA	205	28	0.84	(0.54,	1.33)	19	1.51	(0.89,	2.58)	71	1.13	(0.83,	1.52)	16	1.10	(0.64,	1.89)	
Wald-p			1.00				0.89				1.00				1.00			
RUNX2 (rs12333172)																		
CC/CT	2997	472	1.00			188	1.00			927	1.00			207	1.00			1.00
TT	104	20	1.20	(0.68,	2.10)	11	1.67	(0.83,	3.34)	31	0.90	(0.59,	1.40)	17	**2.25**	**(1.29,**	**3.93)**	
Wald-p			1.00				0.90				1.00				**0.04**			
RUNX2 (rs1200428)																		
CC	1753	297	1.00			111	1.00			523	1.00			138	1.00			1.00
CA	1178	178	0.89	(0.71,	1.11)	75	1.01	(0.73,	1.38)	383	1.10	(0.94,	1.30)	75	0.80	(0.59,	1.07)	
AA	170	17	0.58	(0.33,	1.02)	13	1.21	(0.64,	2.29)	53	1.05	(0.74,	1.48)	11	0.82	(0.43,	1.57)	
p-trend			0.57				1.00				1.00				1.00			
RUNX2 (rs598953)																		
TT	1129	203	1.00			71	1.00			354	1.00			83	1.00			0.62
TA	1489	233	**0.85**	**(0.68,**	**1.07)**	97	1.02	(0.73,	1.42)	456	0.99	(0.84,	1.18)	118	1.09	(0.81,	1.47)	
AA	483	56	**0.61**	**(0.43,**	**0.87)**	31	0.97	(0.61,	1.54)	149	1.02	(0.81,	1.29)	23	0.64	(0.39,	1.04)	
p-trend			**0.07**				1.00				1.00				1.00			
RUNX3 (rs2236850)																		
TT	1041	162	1.00			53	1.00			304	1.00			60	1.00			1.00
TC	1463	233	1.06	(0.83,	1.35)	105	1.45	(1.01,	2.07)	461	1.05	(0.88,	1.26)	110	1.26	(0.91,	1.76)	
CC	590	97	1.15	(0.85,	1.57)	41	1.44	(0.92,	2.24)	193	1.06	(0.85,	1.32)	51	1.43	(0.96,	2.12)	
p-trend			1.00				0.44				1.00				0.42			
RUNX3 (rs9438876)																		
AA	846	130	1.00			49	1.00			240	1.00			55	1.00			1.00

	Pre-menopausal								Post-menopausal								
	Controls [a]	**ER+**				**ER-**				**ER+**				**ER-**			
	N	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**p-int[c]**
AG	1463	249	1.03	(0.79,	1.33)	92	1.01	(0.70,	1.48)	476	1.18	(0.98,	1.42)	103	1.14	(0.81,	1.61)	
GG	792	113	0.93	(0.69,	1.27)	58	1.26	(0.83,	1.92)	242	1.04	(0.84,	1.29)	65	1.29	(0.88,	1.89)	
p-trend			1.00				1.00				1.00				0.78			
RUNX3 (rs7517302)																		
TT/TC	2567	406	1.00			161	1.00			772	1.00			174	1.00			1.00
CC	531	86	1.11	(0.83,	1.47)	38	1.21	(0.82,	1.79)	184	1.12	(0.92,	1.36)	50	1.36	(0.97,	1.91)	
Wald-p			1.00				1.00				0.83				0.42			
RUNX3 (rs906296)																		
CC	1936	276	1.00			125	1.00			570	1.00			132	1.00			1.00
CG/GG	1162	216	**1.33**	**(1.07,**	**1.65)**	74	1.02	(0.74,	1.39)	387	1.11	(0.95,	1.30)	92	1.16	(0.87,	1.54)	
Wald-p			**0.06**				1.00				0.76				0.78			
RUNX3 (rs7551188)																		
CC	750	124	1.00			53	1.00			212	1.00			53	1.00			1.00
CT	1521	260	1.03	(0.79,	1.34)	101	0.94	(0.65,	1.36)	490	1.12	(0.92,	1.36)	111	1.03	(0.73,	1.46)	
TT	826	108	0.81	(0.59,	1.10)	45	0.80	(0.52,	1.23)	256	1.06	(0.85,	1.32)	59	0.97	(0.65,	1.44)	
p-trend			0.72				1.00				1.00				0.87			
RUNX3 (rs6688058)																		
GG/GA	3046	485	1.00			196	1.00			934	1.00			219	1.00			1.00
AA	55	7	0.85	(0.35,	2.04)	3	0.81	(0.24,	2.76)	25	1.48	(0.88,	2.49)	5	1.23	(0.47,	3.17)	
Wald-p			1.00				1.00				0.68				0.78			
RUNX3 (rs11249206)																		
TT	918	167	1.00			71	1.00			276	1.00			63	1.00			1.00
TC	1505	234	0.85	(0.67,	1.09)	93	0.81	(0.58,	1.14)	466	0.99	(0.82,	1.19)	105	1.03	(0.74,	1.44)	
CC	627	85	0.72	(0.53,	0.99)	35	0.73	(0.46,	1.13)	199	1.01	(0.81,	1.27)	53	1.27	(0.86,	1.89)	
p-trend			0.25				0.66				1.00				0.78			

	Pre-menopausal								Post-menopausal								
	Controls [a]	**ER+**				**ER-**				**ER+**				**ER-**			
	N	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**		**p-int[c]**
RUNX3 (rs4478762)																		
GG/GA	3066	487	1.00			197	1.00			939	1.00			221	1.00			1.00
AA	33	5	0.94	(0.33,	2.71)	2	0.84	(0.19,	3.78)	20	**2.08**	**(1.12,**	**3.86)**	3	1.21	(0.36,	4.12)	
Wald-p			1.00				1.00				**0.13**				0.78			
TGF-β1 (rs1800469)																		
CC	1231	190	1.00			74	1.00			383	1.00			73	1.00			0.30
CT	1368	226	1.02	(0.81,	1.30)	84	0.92	(0.65,	1.30)	394	0.98	(0.82,	1.16)	115	1.43	(1.05,	1.95)	
TT	451	65	0.98	(0.69,	1.39)	36	1.27	(0.81,	2.00)	156	1.16	(0.92,	1.46)	34	1.24	(0.80,	1.91)	
p-trend			0.99				0.92				0.37				0.27			
TGF-β1 (rs4803455)																		
CC	898	134	1.00			59	1.00			291	1.00			64	1.00			0.48
CA	1414	200	0.98	(0.75,	1.27)	74	0.85	(0.58,	1.23)	406	**0.83**	**(0.69,**	**1.00)**	92	0.86	(0.62,	1.21)	
AA	560	95	1.14	(0.82,	1.59)	32	0.94	(0.58,	1.53)	162	**0.82**	**(0.65,**	**1.04)**	32	0.76	(0.49,	1.19)	
p-trend			0.99				0.92				**0.13**				0.27			
TGF-βRI (rs6478974)																		
TT/TA	2532	400	1.00			167	1.00			772	1.00			171	1.00			0.98
AA	568	92	1.02	(0.77,	1.35)	32	0.92	(0.61,	1.39)	186	1.04	(0.86,	1.27)	53	1.42	(1.02,	1.98)	
Wald-p			0.89				1.00				1.00				0.16			
TGF-βRI (rs1571590)																		
AA	2193	345	1.00			140	1.00			665	1.00			164	1.00			0.98
AG/GG	907	147	1.11	(0.87,	1.41)	59	1.12	(0.80,	1.57)	294	1.00	(0.85,	1.19)	60	0.86	(0.63,	1.18)	
Wald-p			0.68				1.00				1.00				1.00			
TGF-βRI (rs1013186)																		
GG	2192	344	1.00			140	1.00			665	1.00			164	1.00			0.98
GA/AA	909	148	1.12	(0.89,	1.43)	59	1.12	(0.80,	1.57)	294	1.00	(0.84,	1.18)	60	0.86	(0.63,	1.18)	

	Pre-menopausal								Post-menopausal							
	Controls [a]	**ER+**				**ER-**			**ER+**				**ER-**			
	N	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**	**N**	**OR[b]**	**(95% CI)**		**N**	**OR[b]**	**(95% CI)**	**p-int[c]**
Wald-p			0.68				1.00			1.00				1.00		
TGF-βRI (rs11568785)																
AA	2701	410	1.00			174	1.00		810	1.00			198	1.00		0.98
AG/GG	400	82	1.34	(0.99,	1.81)	25	1.03	(0.65, 1.64)	149	1.20	(0.96,	1.49)	26	0.88	(0.57, 1.36)	
Wald-p			0.19				1.00			0.43				1.00		
TGF-βRI (rs10733710)																
GG/GA	2835	466	1.00			190	1.00		895	1.00			208	1.00		0.98
AA	265	26	**0.58**	**(0.37,**	**0.92)**	9	**0.45**	**(0.22, 0.91)**	64	0.82	(0.61,	1.11)	15	0.78	(0.45, 1.35)	
Wald-p			**0.08**				**0.10**			0.61				1.00		

[a] ER data were compared with 3,214 controls from sites where cases have ER data Mexico data is excluded because they do not have data for ER status (n=1,810)

[b] Odds ratios (OR) and 95% confidence intervals (CI) adjusted for age, study, and genetic admixture; OR in **bold text** indicates significance ($p \leq 0.05$) or suggestive of an association ($p \leq 0.15$) for p-trend (per-allele) or Wald-p after adjustment for multiple comparisons

[c] p-value for interaction term (SNP*menopause) for ER+ or ER- breast cancer as the outcome; Bonferroni-Holms p-value adjustment for multiple comparisons shown in parentheses

Table 18b. TGF-β signaling and ERα SNPs and breast cancer defined by ER status, stratified by Native American ancestry (*Full table*)

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]	ER+				ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
RUNX1 (rs7279383)																	
CC	2249	639	1.00			173	1.00		455	1.00			165	1.00			1.00
CG/GG	965	293	0.88	(0.74,	1.04)	71	0.78	(0.58, 1.04)	124	0.85	(0.67,	1.07)	42	0.80	(0.55,	1.14)	
Wald-p			1.00				0.74			0.97				1.00			
RUNX1 (rs2268288)																	
TT/TC	3122	889	1.00			233	1.00		570	1.00			200	1.00			1.00
CC	93	44	1.25	(0.85,	1.83)	11	1.15	(0.60, 2.22)	9	1.03	(0.47,	2.29)	7	2.39	(0.99,	5.80)	
Wald-p			1.00				1.00			1.00				0.43			
RUNX1 (rs2252585)																	
TT/TC	2822	857	1.00			227	1.00		485	1.00			170	1.00			1.00
CC	393	76	0.99	(0.74,	1.32)	17	0.82	(0.49, 1.38)	94	0.89	(0.68,	1.16)	36	0.97	(0.66,	1.43)	
Wald-p			1.00				1.00			1.00				1.00			
RUNX1 (rs11701453)																	
GG	2179	595	1.00			159	1.00		427	1.00			138	1.00			1.00
GC	931	296	1.03	(0.87,	1.22)	80	1.06	(0.79, 1.41)	146	0.96	(0.76,	1.20)	64	1.30	(0.94,	1.80)	
CC	105	41	1.05	(0.71,	1.56)	5	0.49	(0.19, 1.22)	6	0.51	(0.21,	1.24)	4	1.01	(0.35,	2.93)	
p-trend			1.00				1.00			1.00				1.00			
RUNX1 (rs8127225)																	
TT	2132	682	1.00			174	1.00		276	1.00			116	1.00			0.38
TC/CC	1078	251	1.02	(0.86,	1.22)	70	1.11	(0.82, 1.50)	303	**1.39**	**(1.14,**	**1.69)**	91	0.98	(0.73,	1.32)	
Wald-p			1.00				1.00			**0.008**				1.00			
RUNX1 (rs1474479)																	
GG/GA	2896	811	1.00			216	1.00		555	1.00			199	1.00			1.00
AA	319	122	0.87	(0.69,	1.10)	27	0.74	(0.49, 1.13)	24	1.43	(0.85,	2.41)	8	1.31	(0.60,	2.85)	

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]		ER+			ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
Wald-p			1.00				1.00				0.97				1.00			
RUNX1 (rs1883066)																		
GG	2569	712	1.00			182	1.00			478	1.00			181	1.00			1.00
GC/CC	645	221	1.02	(0.85,	1.23)	62	1.13	(0.83,	1.54)	101	1.13	(0.87,	1.46)	26	0.79	(0.51,	1.22)	
Wald-p			1.00				1.00				1.00				1.00			
RUNX1 (rs7279123)																		
CC	1935	525	1.00			139	1.00			365	1.00			141	1.00			1.00
CT	1083	341	0.96	(0.81,	1.13)	88	0.93	(0.70,	1.24)	190	1.23	(1.00,	1.52)	61	1.02	(0.74,	1.41)	
TT	184	63	0.94	(0.69,	1.30)	15	0.88	(0.50,	1.55)	23	1.09	(0.66,	1.81)	5	0.63	(0.24,	1.60)	
p-trend			1.00				1.00				0.75				1.00			
RUNX2 (rs1321075)																		
CC	1918	635	1.00			175	1.00			245	1.00			88	1.00			0.98
CA	1102	269	1.07	(0.90,	1.28)	59	0.88	(0.64,	1.20)	267	1.13	(0.92,	1.39)	87	1.02	(0.74,	1.40)	
AA	195	29	1.08	(0.68,	1.72)	10	1.32	(0.66,	2.66)	66	1.17	(0.84,	1.62)	32	1.56	(1.00,	2.44)	
p-trend			1.00				0.88				0.77				1.00			
RUNX2 (rs17209895)																		
TT	2002	531	1.00			117	1.00			439	1.00			150	1.00			0.98
TC	1054	347	0.92	(0.78,	1.08)	108	1.29	(0.98,	1.71)	128	0.89	(0.71,	1.13)	52	1.06	(0.76,	1.50)	
CC	159	55	0.89	(0.64,	1.25)	19	1.39	(0.82,	2.36)	12	0.69	(0.36,	1.34)	5	0.87	(0.34,	2.25)	
p-trend			1.00				0.38				0.77				1.00			
RUNX2 (rs2677108)																		
TT	952	320	1.00			77	1.00			132	1.00			49	1.00			0.98
TC	1559	468	1.03	(0.87,	1.23)	120	1.12	(0.83,	1.53)	293	1.06	(0.83,	1.36)	95	0.92	(0.64,	1.34)	
CC	702	145	0.88	(0.69,	1.12)	47	1.21	(0.82,	1.79)	154	0.98	(0.75,	1.30)	63	1.07	(0.72,	1.61)	
p-trend			1.00				0.73				0.88				1.00			

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]		ER+			ER-				ER+			ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
RUNX2 (rs2819854)																		
TT	819	220	1.00			59	1.00			147	1.00			68	1.00			0.60
TC	1599	469	0.99	(0.82,	1.21)	132	1.04	(0.75,	1.44)	294	1.16	(0.91,	1.47)	83	0.71	(0.50,	1.01)	
CC	795	244	1.02	(0.81,	1.28)	53	0.80	(0.54,	1.19)	137	1.10	(0.83,	1.45)	56	0.99	(0.67,	1.45)	
p-trend			1.00				0.73				0.77				1.00			
RUNX2 (rs2790093)																		
AA	1450	394	1.00			110	1.00			290	1.00			104	1.00			0.98
AG	1458	441	1.10	(0.93,	1.30)	109	1.00	(0.75,	1.33)	242	0.85	(0.69,	1.04)	84	0.82	(0.60,	1.12)	
GG	307	98	1.12	(0.85,	1.47)	25	1.07	(0.67,	1.71)	47	0.82	(0.57,	1.19)	19	0.91	(0.54,	1.55)	
p-trend			1.00				0.83				0.77				1.00			
RUNX2 (rs9463090)																		
GG	2051	586	1.00			131	1.00			387	1.00			132	1.00			0.98
GA	1016	304	1.01	(0.85,	1.19)	89	**1.29**	**(0.97,**	**1.72)**	176	1.00	(0.81,	1.24)	58	0.94	(0.68,	1.32)	
AA	144	42	0.82	(0.56,	1.20)	24	**2.03**	**(1.24,**	**3.30)**	15	0.80	(0.44,	1.45)	17	2.73	(1.51,	4.95)	
p-trend			1.00				**0.03**				0.77				0.73			
RUNX2 (rs2396441)																		
CC	785	251	1.00			73	1.00			131	1.00			48	1.00			0.38
CT	1684	460	0.83	(0.69,	1.01)	115	0.72	(0.52,	0.98)	278	**1.02**	**(0.80,**	**1.31)**	108	1.11	(0.77,	1.60)	
TT	745	222	0.88	(0.70,	1.10)	56	0.76	(0.52,	1.10)	170	**1.45**	**(1.10,**	**1.91)**	51	1.19	(0.78,	1.83)	
p-trend			1.00				0.48				**0.08**				1.00			
RUNX2 (rs1316330)																		
GG	1971	504	1.00			141	1.00			414	1.00			149	1.00			0.98
GT	1109	381	1.10	(0.94,	1.30)	87	0.92	(0.69,	1.22)	156	0.89	(0.71,	1.11)	55	0.90	(0.64,	1.25)	
TT	133	48	1.01	(0.70,	1.45)	16	1.24	(0.71,	2.17)	9	0.62	(0.29,	1.31)	3	0.53	(0.16,	1.77)	
p-trend			1.00				0.95				0.77				1.00			

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]		ER+			ER-				ER+				ER-			
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
RUNX2 (rs7750470)																		
TT	2060	586	1.00			147	1.00			389	1.00			126	1.00			0.78
TC	1024	315	1.12	(0.94,	1.32)	81	1.12	(0.84,	1.50)	168	0.85	(0.69,	1.06)	74	1.16	(0.85,	1.59)	
CC	131	32	0.90	(0.59,	1.39)	16	1.74	(0.98,	3.10)	22	0.84	(0.50,	1.38)	7	0.78	(0.35,	1.76)	
p-trend			1.00				0.43				0.77				1.00			
RUNX2 (rs6930053)																		
CC	1280	334	1.00			106	1.00			245	1.00			95	1.00			0.56
CT	1516	447	0.90	(0.76,	1.08)	108	**0.69**	**(0.52,**	**0.93)**	269	1.22	(0.99,	1.50)	93	1.10	(0.81,	1.49)	
TT	419	151	1.00	(0.79,	1.27)	30	**0.63**	**(0.41,**	**0.97)**	65	1.31	(0.94,	1.83)	19	0.98	(0.58,	1.67)	
p-trend			1.00				**0.09**				0.37				1.00			
RUNX2 (rs12208240)																		
GG	2599	789	1.00			204	1.00			420	1.00			160	1.00			0.98
GA	583	136	0.94	(0.75,	1.17)	37	0.97	(0.66,	1.40)	147	1.22	(0.97,	1.53)	43	0.94	(0.65,	1.35)	
AA	33	7	1.02	(0.40,	2.57)	3	1.86	(0.52,	6.66)	12	1.44	(0.70,	2.99)	2	0.65	(0.15,	2.82)	
p-trend			1.00				0.83				0.53				1.00			
RUNX2 (rs12209785)																		
AA/AG	3024	876	1.00			224	1.00			540	1.00			190	1.00			0.98
GG	190	56	1.19	(0.85,	1.68)	20	1.70	(1.02,	2.82)	39	0.94	(0.64,	1.39)	17	1.19	(0.69,	2.04)	
Wald-p			1.00				0.33				0.77				1.00			
RUNX2 (rs10948238)																		
CC/CT	2742	775	1.00			191	1.00			496	1.00			173	1.00			0.78
TT	470	155	1.20	(0.96,	1.49)	53	**1.68**	**(1.20,**	**2.34)**	83	0.94	(0.72,	1.25)	34	1.11	(0.74,	1.65)	
Wald-p			1.00				**0.03**				0.77				1.00			
RUNX2 (rs13201287)																		
GG/GA	3003	870	1.00			224	1.00			536	1.00			187	1.00			0.98

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]		ER+			ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
AA	212	63	1.24	(0.89,	1.71)	20	1.58	(0.96, 2.61)	43	0.90	(0.62,	1.30)	20	1.23	(0.74,	2.03)	
Wald-p			1.00				0.43			0.77				1.00			
RUNX2 (rs12333172)																	
CC/CT	3109	892	1.00			226	1.00		568	1.00			195	1.00			0.98
TT	106	40	1.03	(0.70,	1.54)	18	1.77	(1.03, 3.04)	11	0.82	(0.41,	1.65)	12	**2.58**	**(1.30,**	**5.13)**	
Wald-p			1.00				0.33			0.77				**0.07**			
RUNX2 (rs1200428)																	
CC	1818	530	1.00			150	1.00		320	1.00			112	1.00			0.33
CA	1220	359	1.10	(0.93,	1.30)	76	0.84	(0.62, 1.12)	228	0.95	(0.77,	1.16)	84	0.99	(0.73,	1.35)	
AA	177	44	1.19	(0.81,	1.75)	18	1.83	(1.06, 3.16)	31	0.66	(0.43,	1.01)	11	0.67	(0.35,	1.30)	
p-trend			1.00				0.73			0.77				1.00			
RUNX2 (rs598953)																	
TT	1175	378	1.00			94	1.00		199	1.00			68	1.00			0.41
TA	1539	421	0.89	(0.75,	1.06)	123	1.06	(0.80, 1.41)	298	1.06	(0.86,	1.32)	105	1.12	(0.81,	1.56)	
AA	501	134	1.07	(0.83,	1.36)	27	0.89	(0.57, 1.41)	82	0.71	(0.53,	0.96)	34	0.85	(0.55,	1.32)	
p-trend			1.00				0.84			0.67				1.00			
RUNX3 (rs2236850)																	
TT	1073	296	1.00			63	1.00		182	1.00			53	1.00			1.00
TC	1526	451	1.03	(0.86,	1.24)	126	1.31	(0.95, 1.81)	281	1.13	(0.90,	1.41)	105	**1.47**	**(1.03,**	**2.09)**	
CC	609	185	1.00	(0.80,	1.25)	54	1.32	(0.89, 1.94)	116	1.26	(0.95,	1.67)	47	**1.79**	**(1.17,**	**2.74)**	
p-trend			0.96				0.55			0.49				**0.03**			
RUNX3 (rs9438876)																	
AA	871	215	1.00			57	1.00		173	1.00			51	1.00			1.00
AG	1527	464	1.04	(0.86,	1.28)	108	0.88	(0.62, 1.24)	290	1.14	(0.91,	1.43)	103	**1.38**	**(0.97,**	**1.98)**	
GG	817	253	0.91	(0.73,	1.14)	79	1.04	(0.72, 1.50)	116	1.11	(0.84,	1.47)	52	**1.70**	**(1.12,**	**2.57)**	

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]		ER+			ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
p-trend			0.72				1.00			1.00				**0.06**			
RUNX3 (rs7517302)																	
TT/TC	2664	744	1.00			185	1.00		482	1.00			171	1.00			1.00
CC	548	187	1.14	(0.93,	1.39)	59	1.43	(1.04, 1.97)	96	1.10	(0.84,	1.43)	36	1.18	(0.80,	1.74)	
Wald-p			0.72				0.16			1.00				1.00			
RUNX3 (rs906296)																	
CC	2009	535	1.00			140	1.00		347	1.00			134	1.00			1.00
CG/GG	1203	396	1.10	(0.94,	1.29)	104	1.08	(0.83, 1.42)	232	**1.33**	**(1.08,**	**1.62)**	73	1.10	(0.81,	1.50)	
Wald-p			0.72				1.00			**0.04**				1.00			
RUNX3 (rs7551188)																	
CC	781	198	1.00			63	1.00		148	1.00			49	1.00			1.00
CT	1569	479	1.03	(0.84,	1.26)	120	0.83	(0.59, 1.15)	300	1.21	(0.96,	1.54)	102	1.23	(0.85,	1.77)	
TT	860	256	0.97	(0.77,	1.21)	61	0.72	(0.49, 1.05)	130	1.01	(0.77,	1.34)	53	1.26	(0.83,	1.91)	
p-trend			0.72				0.45			1.00				1.00			
RUNX3 (rs6688058)																	
GG/GA	3155	910	1.00			242	1.00		568	1.00			200	1.00			0.55
AA	60	23	1.71	(0.97,	2.99)	2	0.58	(0.14, 2.46)	11	0.79	(0.40,	1.59)	7	1.41	(0.61,	3.25)	
Wald-p			0.31				1.00			1.00				1.00			
RUNX3 (rs11249206)																	
TT	951	246	1.00			65	1.00		211	1.00			81	1.00			1.00
TC	1560	452	0.83	(0.69,	1.01)	116	0.83	(0.60, 1.15)	279	1.12	(0.91,	1.39)	91	0.95	(0.69,	1.32)	
CC	649	219	0.86	(0.68,	1.08)	60	0.90	(0.62, 1.32)	77	0.95	(0.70,	1.29)	35	1.12	(0.73,	1.72)	
p-trend			0.72				1.00			1.00				1.00			
RUNX3 (rs4478762)																	
GG/GA	3178	915	1.00			242	1.00		570	1.00			203	1.00			1.00

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]		ER+			ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
AA	35	18	**2.45**	**(1.23,**	**4.91)**	2	1.00	(0.22, 4.43)	9	1.04	(0.47,	2.29)	4	1.30	(0.44,	3.87)	
Wald-p			**0.07**				1.00			1.00				1.00			
TGF-β1 (rs1800469)																	
CC	1269	432	1.00			102	1.00		160	1.00			57	1.00			0.42
CT	1425	372	0.97	(0.82,	1.15)	108	1.18	(0.89, 1.58)	278	1.06	(0.84,	1.34)	104	1.08	(0.76,	1.53)	
TT	469	100	0.89	(0.68,	1.16)	30	1.15	(0.74, 1.78)	133	**1.37**	**(1.04,**	**1.81)**	43	1.20	(0.78,	1.85)	
p-trend			0.84				0.55			**0.07**				0.81			
TGF-β1 (rs4803455)																	
CC	927	240	1.00			67	1.00		198	1.00			59	1.00			0.42
CA/AA	2041	646	0.97	(0.81,	1.17)	157	0.84	(0.62, 1.14)	250	**0.79**	**(0.64,**	**0.99)**	92	1.02	(0.72,	1.45)	
Wald-p			0.84				0.55			**0.07**				0.92			
TGF-βR1 (rs6478974)																	
TT/TA	2634	714	1.00			186	1.00		510	1.00			174	1.00			1.00
AA	580	218	1.12	(0.92,	1.35)	58	1.16	(0.84, 1.59)	69	0.90	(0.67,	1.21)	33	1.30	(0.87,	1.95)	
Wald-p			1.00				1.00			1.00				0.74			
TGF-βR1 (rs1571590)																	
AA	2277	601	1.00			161	1.00		452	1.00			163	1.00			1.00
AG	859	306	1.05	(0.89,	1.25)	75	0.97	(0.72, 1.30)	121	1.04	(0.82,	1.32)	40	0.96	(0.66,	1.39)	
GG	78	26	0.76	(0.48,	1.22)	8	0.91	(0.43, 1.94)	6	1.44	(0.52,	4.00)	4	2.70	(0.83,	8.79)	
p-trend			1.00				1.00			1.00				1.00			
TGF-βR1 (rs1013186)																	
GG	2275	601	1.00			161	1.00		451	1.00			163	1.00			1.00
GA	862	306	1.05	(0.88,	1.24)	75	0.96	(0.72, 1.29)	122	1.05	(0.83,	1.34)	40	0.96	(0.66,	1.39)	
AA	78	26	0.77	(0.48,	1.22)	8	0.92	(0.43, 1.95)	6	1.44	(0.52,	4.01)	4	2.70	(0.83,	8.79)	
p-trend			1.00				1.00			1.00				1.00			

	Low (0-28%)								Moderate to High (29-100%)								
	Controls[a]	ER+				ER-			ER+				ER-				
	N	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)	N	OR[b]	(95% CI)		N	OR[b]	(95% CI)		p-int[c]
TGF-βR1 (rs11568785)																	
AA/AG	3199	923	1.00			244	1.00		578	1.00			205	1.00			1.00
GG	16	10	1.37	(0.61,	3.08)	0	0.00	(0.00, 0.00)	1	2.48	(0.15,	39.9)	2	10.96	(0.98,	123.18)	
Wald-p			1.00				1.00			1.00				0.24			
TGF-βR1 (rs10733710)																	
GG/GA	2938	890	1.00			229	1.00		527	1.00			192	1.00			0.18
AA	276	43	0.90	(0.62,	1.31)	14	1.20	(0.67, 2.15)	52	**0.64**	**(0.46,**	**0.89)**	15	**0.48**	**(0.28,**	**0.84)**	
Wald-p			1.00				1.00			**0.04**				**0.05**			

[a] ER data were compared with 3,214 controls from sites where cases have ER data Mexico data is excluded because they do not have data for ER status (n=1,810)

[b] Odds ratios (OR) and 95% confidence intervals (CI) adjusted for age, study, and genetic admixture; OR in **bold text** indicates significance ($p \leq 0.05$) or suggestive of an association ($p \leq 0.15$) for p-trend (per-allele) or Wald-p after adjustment for multiple comparisons

[c] p-value for interaction term (SNP*admixture) for ER+ or ER- breast cancer as the outcome; Bonferroni-Holms p-value adjustment for multiple comparisons shown

Table 22b. SNP-SNP interactions between ERα and TGF-β signaling genes (*Full table*)

Gene (SNP) [a]	Gene (SNP) [a]	Cases/Controls	Combined Risk		
			OR (95% CI) [b]	Wald-p [c]	p-int [d]
ERα (rs1801132)	*TGF-β1* (rs4803455)			0.02	0.11
CC/CG	CC	202/198	1.00 (REF)		
CC/CG	CA/AA	428/474	0.89 (0.70-1.13)		
GG	**CC**	**22/7**	**3.12 (1.30-7.48)**		
GG	CA/AA	31/26	1.19 (0.68-2.08)		
ERα (rs3798577)	*TGF-β1* (rs4803455)			0.06	0.59
TT/TC	CC	173/165	1.00 (REF)		
TT/TC	CA/AA	353/412	0.82 (0.64-1.06)		
CC	CC	51/40	1.21 (0.76-1.94)		
CC	CA/AA	106/88	1.17 (0.81-1.67)		
ERα (rs1801132)	*TGF-βR1* (rs6478974)			0.06	0.21
CC/CG	TT/TA	497/537	1.00 (REF)		
CC/CG	AA	133/135	1.07 (0.82-1.40)		
GG	TT/TA	42/30	1.53 (0.94-2.49)		
GG	**AA**	**11/3**	**4.01 (1.11-14.47)**		
ERα (rs3798577)	*TGF-βR1* (rs6478974)			0.08	0.30
TT/TC	TT/TA	413/468	1.00 (REF)		
TT/TC	AA	113/109	1.18 (0.89-1.59)		
CC	**TT/TA**	**126/99**	**1.45 (1.08-1.95)**		
CC	AA	31/29	1.22 (0.72-2.07)		
ERα (rs1801132)	*RUNX1* (rs7279383)			0.06	0.18
CC/C**G**	CC	437/450	1.00 (REF)		
CC/C**G**	C**G**/**GG**	193/222	0.90 (0.71-1.24)		
GG	CC	32/24	1.38 (0.80-2.38)		
GG	C**G**/**GG**	21/9	2.42 (1.10-5.36)		
ERα (rs1801132)	*RUNX1* (rs2268288)			0.19	0.98
CC/C**G**	TT/T**C**	604/644	1.00 (REF)		
CC/C**G**	**CC**	26/28	0.98 (0.57-1.70)		
GG	TT/T**C**	51/33	1.65 (1.05-2.60)		
GG	**CC**	2/0	--		
ERα (rs1801132)	*RUNX1* (rs8127225)			0.11	0.50
CC/C**G**	TT	440/474	1.00 (REF)		
CC/C**G**	T**C**/**CC**	190/198	1.04 (0.82-1.33)		
GG	TT	37/21	1.91 (1.10-3.33)		
GG	T**C**/**CC**	16/12	1.44 (0.67-3.09)		
ERα (rs1801132)	*RUNX1* (rs7279123)			0.09	0.91
CC/C**G**	CC	386/392	1.00 (REF)		
CC/C**G**	C**T**/**TT**	242/276	0.88 (0.70-1.10)		
GG	CC	33/20	1.68 (0.95-2.99)		
GG	C**T**/**TT**	20/13	1.56 (0.77-3.19)		

Gene (SNP) [a]	Gene (SNP) [a]	Cases/Controls	Combined Risk		
			OR (95% CI) [b]	Wald-p [c]	p-int [d]
ERα (rs3798577)	*RUNX1* (rs7279383)			0.14	0.70
TT/T**C**	CC	363/387	1.00 (REF)		
TT/T**C**	C**G**/**GG**	163/190	0.92 (0.71-1.19)		
CC	CC	106/87	1.31 (0.95-1.80)		
CC	C**G**/**GG**	51/41	1.34 (0.87-2.08)		
ERα (rs3798577)	*RUNX1* (rs2268288)			0.16	0.91
TT/T**C**	TT/T**C**	505/555	1.00 (REF)		
TT/T**C**	**CC**	21/22	1.04 (0.56-1.92)		
CC	TT/T**C**	150/122	1.36 (1.04-1.78)		
CC	**CC**	7/6	1.31 (0.44-3.94)		
ERα (rs3798577)	*RUNX1* (rs8127225)			0.13	0.46
TT/T**C**	TT	364/398	1.00 (REF)		
TT/T**C**	T**C**/**CC**	162/179	0.99 (0.76-1.28)		
CC	TT	113/97	1.28 (0.94-1.74)		
CC	T**C**/**CC**	44/31	1.58 (0.97-2.56)		
ERα (rs3798577)	*RUNX1* (rs7279123)			0.14	0.97
TT/T**C**	CC	321/335	1.00 (REF)		
TT/T**C**	C**T**/**TT**	205/238	0.89 (0.70-1.14)		
CC	CC	98/77	1.33 (0.95-1.86)		
CC	C**T**/**TT**	57/51	1.17 (0.78-1.77)		
ERα (rs1801132)	*RUNX2* (rs12209785)			0.04	0.51
CC/C**G**	AA/A**G**	587/640	1.00 (REF)		
CC/C**G**	**GG**	43/32	1.47 (0.91-2.35)		
GG	AA/A**G**	49/30	1.78 (1.12-2.85)		
GG	**GG**	4/3	1.51 (0.34-6.81)		
ERα (rs1801132)	*RUNX2* (rs10948238)			0.11	0.45
CC/C**G**	CC/C**T**	527/562	1.00 (REF)		
CC/C**G**	**TT**	103/110	1.00 (0.74-1.34)		
GG	CC/C**T**	44/25	1.88 (1.13-3.12)		
GG	**TT**	9/8	1.22 (0.47-3.19)		
ERα (rs1801132)	*RUNX2* (rs13201287)			0.03	0.31
CC/C**G**	GG/G**A**	583/637	1.00 (REF)		
CC/C**G**	**AA**	47/35	1.47 (0.94-2.31)		
GG	GG/G**A**	50/30	1.83 (1.15-2.91)		
GG	**AA**	3/3	1.11 (0.22-5.51)		
ERα (rs3798577)	*RUNX2* (rs12209785)			0.06	0.77
TT/T**C**	AA/A**G**	492/549	1.00 (REF)		
TT/T**C**	**GG**	34/28	1.35 (0.81-2.27)		
CC	AA/A**G**	144/121	1.34 (1.02-1.75)		
CC	**GG**	13/7	2.13 (0.84-5.40)		
ERα (rs3798577)	*RUNX2* (rs10948238)			0.15	0.70
TT/T**C**	CC/C**T**	438/481	1.00 (REF)		

Gene (SNP) [a]	Gene (SNP) [a]	Cases/Controls	Combined Risk		
			OR (95% CI) [b]	Wald-p [c]	p-int [d]
TT/T**C**	**TT**	88/96	1.01 (0.73-1.38)		
CC	CC/C**T**	133/103	1.39 (1.04-1.85)		
CC	**TT**	24/22	1.21 (0.67-2.20)		
ERα (rs3798577)	*RUNX2* (rs13201287)			0.06	0.75
TT/T**C**	GG/G**A**	489/548	1.00 (REF)		
TT/T**C**	**AA**	37/29	1.43 (0.87-2.63)		
CC	GG/G**A**	144/119	1.36 (1.04-1.79)		
CC	**AA**	13/9	1.65 (0.70-3.91)		
ERα (rs1801132)	*RUNX3* (rs7517302)			0.05	0.34
CC/C**G**	TT/T**C**	502/552	1.00 (REF)		
CC/C**G**	**CC**	128/120	1.17 (0.89-1.54)		
GG	TT/T**C**	41/29	1.56 (0.95-2.54)		
GG	**CC**	12/4	3.37 (1.08-10.53)		
ERα (rs1801132)	*RUNX3* (rs906296)			0.03	0.56
CC/C**G**	CC	364/421	1.00 (REF)		
CC/C**G**	C**G**/**GG**	266/251	1.23 (0.98-1.53)		
GG	CC	28/17	1.92 (1.03-3.57)		
GG	C**G**/**GG**	25/16	1.80 (0.95-3.43)		
ERα (rs1801132)	*RUNX3* (rs4478762)			0.06	--
CC/C**G**	GG/G**A**	621/661	1.00 (REF)		
CC/C**G**	**AA**	9/11	0.87 (0.36-2.12)		
GG	GG/G**A**	53/33	1.72 (1.10-2.69)		
GG	**AA**	0/0	--		
ERα (rs3798577)	*RUNX3* (rs7517302)			0.08	0.92
TT/T**C**	TT/T**C**	421/477	1.00 (REF)		
TT/T**C**	**CC**	105/100	1.19 (0.88-1.61)		
CC	TT/T**C**	122/104	1.34 (1.00-1.80)		
CC	**CC**	35/24	1.65 (0.96-2.83)		
ERα (rs3798577)	*RUNX3* (rs906296)			0.03	0.65
TT/T**C**	CC	299/359	1.00 (REF)		
TT/T**C**	C**G**/**GG**	227/218	1.25 (0.98-1.60)		
CC	CC	93/79	1.43 (1.02-2.00)		
CC	C**G**/**GG**	64/49	1.58 (1.05-2.37)		
ERα (rs3798577)	*RUNX3* (rs4478762)			0.15	0.96
TT/T**C**	GG/G**A**	519/568	1.00 (REF)		
TT/T**C**	**AA**	7/9	0.85 (0.31-2.30)		
CC	GG/G**A**	155/126	1.36 (1.04-1.77)		
CC	**AA**	2/2	1.09 (0.15-7.76)		

[a]Minor alleles are denoted in **bold text**
[b]OR and 95% CI adjusted for age and genetic admixture
[c]Wald-p for model; [d] Interaction p-value for interaction term in model (SNP*SNP)

APPENDIX C:

APPROVAL LETTERS FROM THE INSTITUTIONAL REVIEW BOARD

UNIVERSITY OF LOUISVILLE
It's Happening Here.

Human Subjects Protection Program Office
MedCenter One – Suite 200
501 E. Broadway
Louisville, KY 40202-1798
Office: 502.852.5188 Fax: 502.852.2164

March 13, 2008

Kathy Baumgartner, Ph.D.
Epidemiology & Population Health
K Building, Room 4056
555 South Floyd Street
Louisville, Kentucky 40202

RE: 238.05 - 4-Corners Women's Health Study

Dear Doctor Baumgartner:

The following items have been received by the Human Subjects Protection Program Office and approved by the chair of the Institutional Review Board (IRB) through the expedited review procedure according to 45 CFR 46.110(B):

- Amendment #11, dated 3/11/2008

The modifications include:

- The addition of the following personnel to the study: Avonne Connor, MPH and Stephanie Denkhoff, BS.
- The removal of Laura Herman from the study.

The committee will be advised of this action at their next full board meeting. Please send all inquires and electronic revised/requested items to our office email address at hsppofc@louisville.edu.

Sincerely,

Frank A. Walker, M.D., Vice Chair,
Biomedical Institutional Review Board

FAW/elp

UNIVERSITY OF LOUISVILLE
It's Happening Here.

Human Subjects Protection Program Office
MedCenter One – Suite 200
501 E. Broadway
Louisville, KY 40202-1798
Office: 502.852.5188 Fax: 502.852.2164

Friday, July 13, 2012

Kathy Baumgartner, Ph.D.
485 E Gray Street
Room 230
Louisville, KY 40202

RE: IRB# 238.05/4-Corners Women's Health Study

Dear Dr. Baumgartner:

The following items have been received by the Human Subjects Protection Program Office and approved by the chair of the Institutional Review Board (IRB) through the expedited review procedure according to 45 CFR 46.110(B):

- Amendment #23, dated 07/10/2012
- Brief Dissertation Proposal, dated 07/10/2012

The modifications include:

- Submission of Stephanie Denkhoff's brief proposal detailing the specific aims for the research being conducted.

The committee will be advised of this action at their next full board meeting. Please send all inquires and electronic revised/requested items to our office email address at hsppofc@louisville.edu

Sincerely,

Julie L. Goldman, MD
Vice Chair, Biomedical Institutional Review Board

JLG/SLB

CURRICULUM VITAE
STEPHANIE DENKHOFF BOONE

1519 Altawood Dr. ♦ Clarksville, Indiana 47129 ♦ (502) 648-4353 ♦ srdenk01@louisville.edu

EDUCATION

SCHOOL OF PUBLIC HEALTH AND INFORMATION SCIENCES,
UNIVERSITY OF LOUISVILLE
Doctor of Philosophy, Epidemiology and Population Health
Specialization: Breast Cancer Research
Graduation date: May, 2013

SCHOOL OF PUBLIC HEALTH AND INFORMATION SCIENCES,
UNIVERSITY OF LOUISVILLE
Master in Public Health, with a concentration in Epidemiology
Practicum Experience: Leukemia and Lymphoma Society, Louisville Chapter
Graduation date: May, 2008.

SPALDING UNIVERSITY, Louisville, Kentucky
Bachelor of Science, Major: Biology, Minor: Chemistry; cum laude
Graduation date: May, 2006

AWARDED GRANTS, SCHOLARSHIPS, AND HONORS

♦ 2013	Dean's Citation Award, University of Louisville
♦ 2013	Dissertation Completion Award, University of Louisville
♦ 2012	Sponsored Research Tuition Award, University of Louisville
♦ 2011	1st place, Public Health Student, Research! Louisville, University of Louisville Health Sciences Center for poster presentation, "Patterns of All-Cause Mortality Over 10 years by Breast Cancer Tumor Phenotype and Hispanic versus non-Hispanic White Ethnicity"
♦ 2010-present	Susan B. Komen, Dissertation Training Grant, University of Louisville
♦ 2009-present	Graduate Research Assistantship, University of Louisville
♦ 2003-2006	Caritas Award, Spalding University
♦ 2004-2006	Larry Hamfeldt Scholarship, Spalding University Carlton Froess Award, Spalding University
♦ 2002-2006	Honors List (Minimum 3.5 GPA), Spalding University
♦ 2003-2006	Kentucky Intercollegiate Athletic Conference All-Academic Team, Spalding University

EXPERIENCE

UNIVERSITY OF LOUISVILLE, SCHOOL OF PUBLIC HEALTH AND INFORMATION SCIENCES
Graduate Research Assistant, August 2009-present

Assisting in conducting the research study: New Mexico Women's Health Study: *Long Term Quality of Life Phase*. Responsibilities include: interviewing participants, data entry and tracking, and assistance with design and development of materials useful to study, as well as preparation of final documents and manuscripts for submission to National Institute of Health.

UNIVERSITY OF LOUISVILLE, SCHOOL OF PUBLIC HEALTH AND INFORMATION SCIENCES

Program Coordinator, Sr., May 2008 – August 2009

Responsibilities included overseeing all aspects, and coordinating the research study The New Mexico Women's Health Study: *Long Term Quality of Life Phase* funded by the National Institute of Health including; staff and data management. There was thorough involvement with the design and development of materials useful to study through the use of databases such as Oracle, Access, SAS, and graphics software. Develop, implement, and coordinate data collection efforts with staff, as well as collaborate with outside agencies such as a tumor registry and the National Death Index. Identify resources, collect data, develop protocols and data dictionaries, enter and analyze data, and assist in preparing summaries and reports of findings. Work with program faculty and staff to monitor progress on the established project goals and objectives. Coordinate submissions to Institutional Review Board.

UNIVERSITY OF LOUISVILLE, SCHOOL OF PUBLIC HEALTH AND INFORMATION SCIENCES

Research Assistant, September 2007-May 2008, New Mexico Women's Health Study: Long Term Quality of Life Phase. Responsibilities included: interviewing participants, data entry and tracking, and assistance with design and development of materials useful to study.

RELATED EXPERIENCE

- December 2009-February 2011 Assist Dr. Katherine B. Baumgartner, University of Louisville, by preparing evidence tables, summarizing information, assist with rechecking table entries, as well as, the STATA data for the analyses and references for the sections on active and passive smoking and breast cancer for the 2011 Surgeon General's Report (2004-2006 Update) (*The Health Consequences of Smoking and Involuntary Exposure to Tobacco Smoke: An Update)*

SERVICE ACTIVITIES

- August 2008-August 2009 Member, Master in Public Health Advisory Committee University of Louisville, School of Public and Information Sciences.
- May 2007-May 2008 Treasurer, Student Government Association, University of Louisville, School of Public and Information Sciences.
- March 2009-November 2010 Volunteer, Health Career Expo Presenter
    - *Description: provide interactive "hands-on" presentations promoting health professions to middle and high school youth*

PRESENTATIONS

- Stephanie Boone, MPH. Oral Presentation: **TGF-B Signaling Pathway, ERα and the Heterogeneity of Breast Cancer Risk among Hispanic and non-Hispanic White Women.**
    - Susan G. Komen for the Cure Post-Baccalaureate Training in Disparities 2nd Annual Meeting October 26-30, 2012
- Stephanie Denkhoff, MPH, Richard N. Baumgartner, PhD, Kathy B. Baumgartner, PhD. **Ethnicity As a Predictor of Long-Term Quality of Life In Breast Cancer Survivors**
  Abstract/Poster presented at:

- Cancer Survivorship Conference: Cancer Survivorship Research: Translating Science to Care June 14-16, 2012. Arlington, VA.
- Research! Louisville, University of Louisville, September 18, 2012

- Stephanie Denkhoff, MPH, Kathy B. Baumgartner, PhD, Christina Pinkston, MS, Dongyan Yang, MS, M.D. , Richard N. Baumgartner, PhD. **Patterns of All-Cause Mortality Over 10 years by Breast Cancer Tumor Phenotype and Hispanic versus non-Hispanic White Ethnicity.** Abstract/Poster presented at:
  - Research! Louisville, University of Louisville, October 13, 2011
  - James Graham Brown Cancer Retreat, University of Louisville, November 4, 2011
  - American Association of Cancer Research Conference: **The Science of Cancer Health Disparities in Racial/Ethnic Minorities and the Medically Underserved.** September 18-21, 2011. Washington D.C.
- Stephanie Denkhoff, MPH, Kathy B. Baumgartner, PhD, Christina Pinkston, MS, Dongyan Yang, MS, M.D. , Richard N. Baumgartner, PhD. **Hormone Receptor Status and Breast Cancer Survival among Hispanic and non-Hispanic White Women over 10 Years of Follow-Up.** Abstract/Poster presented at: San Antonio Breast Cancer Symposium. December 6-9, 2011. San-Antonio, TX
- Stephanie Denkhoff, MPH, Dongyan Yang, MS, MD, Christina Pinkston, MS, Richard Baumgartner, PhD, Kathy Baumgartner, MA, MS, PhD. **Predictors of Quality of Life among Hispanic and Non-Hispanic White Women: A 15 Year Follow-up Study of Long-Term Breast Cancer Survivors.** Abstract/Poster presented at:
  - Research! Louisville, University of Louisville, October 12, 2010
  - James Graham Brown Cancer Retreat, University of Louisville, November 5, 2010
  - American Association of Cancer Research 102nd Annual Meeting 2011, April 2-6, 2011. Orlando, Florida
- Stephanie Denkhoff, MPH and Carlton Hornung, PhD. **A Review: The Gene-Environment Interaction: Cigarette Smoking, Apolipoprotein E Genotype, and the Risk of Cardiovascular Disease.** Presentation for Journal Club at University of Louisville, January 29, 2010.

PROFESSIONAL MEMBERSHIPS

- Society for Epidemiology Research (SER)
  - Student Member 2009-present
- American Association of Cancer Research (AACR)
  - Associate Member 2010-present
  - Women in Cancer Research (*sub-group, AACR*) 2011-present

PUBLICATIONS

1. Boone, SD, Baumgartner, KB, Joste, NE, Pinkston, CM, Yang, D, Baumgartner RN. (2013) ***The joint contribution of tumor phenotype and education to breast cancer survival disparity between Hispanic and non-Hispanic white women*** [*Submitted, Journal of Clinical Oncology*]

2. Boone, SD, Baumgartner RN, Brock, G, Groves, F, Kerber, R, Pinkston, CM, Connor, AE, A, Hines, L, John, EM, Slattery, ML, Torres-Mejia,G, Wolff, R, Baumgartner, KB (2013) ***TGF-B Signaling Pathway, ERα and the Heterogeneity of Breast Cancer Risk among Hispanic and non-Hispanic White Women.*** [*in preparation*]

3. Boone, SD, Baumgartner, RN, Pinkston, CM, Yang, D, Baumgartner, KB (2013) ***Predictors of Quality of Life among Hispanic and Non-Hispanic White Women: A 15 Year Follow-up Study of Long-Term Breast Cancer Survivors*** [*in preparation*]